Journal of Experimental Psychology

ARTHUR W. MELTON, Editor
Department of Psychology, University of Michigan
Ann Arbor, Michigan

DAVID A. GRANT, Associate Editor
University of Wisconsin

WILLIAM K. ESTES, Associate Editor
Indiana University

Consulting Editors

NORMAN H. ANDERSON, *University of California, Los Angeles*
E. JAMES ARCHER, *University of Wisconsin*
FRED ATTNEAVE, III, *University of Oregon*
CLETUS J. BURKE, *Indiana University*
JAMES DEESE, *Johns Hopkins University*
PAUL M. FITTS, *University of Michigan*
FRANK A. GELDARD, *University of Virginia*
JAMES J. GIBSON, *Cornell University*
HAROLD W. HAKE, *University of Illinois*
ARTHUR L. IRION, *Tulane University*
HOWARD H. KENDLER, *New York University*
HERSCHEL W. LEIBOWITZ, *International Business Machines Corporation, Bethesda, Maryland*
DONALD B. LINDSLEY, *University of California, Los Angeles*
KENNETH MACCORQUODALE, *University of Minnesota*
QUINN MCNEMAR, *Stanford University*
LEO POSTMAN, *University of California, Berkeley*
L. STARLING REID, *University of Virginia*
KENNETH W. SPENCE, *State University of Iowa*
DELOS D. WICKENS, *Ohio State University*

CONTENTS

Effects of Pattern of Reinforcement and Verbal Information on Acquisition, Extinction, and Spontaneous Recovery of the Eyelid CR: T. F. HARTMAN AND D. A. GRANT 217
Pattern Recognition over Distortions, by Human Subjects and by a Computer Simulation of a Model for Human Form Perception: L. UHR, C. VOSSLER, AND J. ULEMAN 227
On the Heterogeneity of Stimulus and Response Elements in the Processing of Information: W. C. HOWELL 235
Space Perception among Unilaterally Paralyzed Children and Adolescents: H. T. BLANE 244
Performance in Instrumental Conditioning as a Joint Function of Time of Deprivation and Sucrose Concentration: J. R. STABLER 248
Concept Identification of Auditory Stimuli as a Function of Amount of Relevant and Irrelevant Information: R. G. BULGARELLA AND E. J. ARCHER 254
Transfer from Verbal-Discrimination to Paired-Associate Learning: W. F. BATTIG, J. M. WILLIAMS, AND J. G. WILLIAMS 258
Relation between Intersession Interval Frequency of Competing Responses and Rate of Learning: D. PREMACK AND R. T. PUTNEY 269
Skin Conductance Levels and Verbal Recall: R. N. BERRY 275
The Effect of Stimulus Similarity on the Acquisition and Extinction of a Conditioned Response: D. P. HUNT 278
Evidence for Shift Effects in the Consummatory Response: D. PREMACK AND W. A. HILLIX .. 284
Reaction Time as a Measure of Intersensory Facilitation: M. HERSHENSON 289
Visual Perception of the Horizontal during Prolonged Exposure to Radial Acceleration on a Centrifuge: B. CLARK AND A. GRAYBIEL 294
Context Factors in Paired-Associate Learning and Recall: D. M. SUNDLAND AND D. D. WICKENS 302
Tests of Three Hypotheses about the Effective Stimulus in Serial Learning: R. K. YOUNG 307
Verbal Concept Attainment: A Function of the Number of Positive and Negative Instances Presented: M. S. MAYZNER 314
Supplementary Report: Monetary Incentive and Motivation in Discrimination Learning—Sex Differences: B. W. ESTES, L. B. MILLER, AND M. E. CURTIN 320

American Psychological Association

Vol. 63, No. 3 March 1962

The JOURNAL OF EXPERIMENTAL PSYCHOLOGY is published monthly, two volumes per year, by the American Psychological Association, Inc., at Prince and Lemon Sts., Lancaster, Pa. The subscription rate per volume is $10.00, or $20.00 annually. Single copies are $2.00. Subscriptions, orders, address changes, and business communications should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St. N.W., Washington 6, D. C.

This Journal publishes original experimental investigations which are intended to contribute toward the development of psychology as an experimental science. Studies with normal human subjects are favored over studies involving abnormal or animal subjects, except when the latter are specifically oriented toward the extension of behavior theory. Experimental psychometric studies and studies in applied experimental psychology or engineering psychology may be accepted if they have broad implications for experimental and theoretical psychology.

Normally, articles exceeding 20 printed pages cannot be accepted. Within this limit, the piecemeal experiment-by-experiment reporting of psychological research is discouraged, and the reporting of the data and conclusions of substantial segments of research efforts is encouraged. Specifically, an integrated series of studies accomplished simultaneously (e.g., most doctoral dissertations) *must* be presented in a single article. Special provision is made for the publication of Supplementary Reports and Replication Reports in articles of not more than 1.5 printed pages (see Editor's Note, this Journal, 1957, **53**, 1–2).

Manuscripts *must* adhere to the conventions concerning reference citations, preparation of tables and figures, manuscript format, etc., as described in the *Publication Manual of the American Psychological Association*. When in doubt about practices of the Journal, authors should examine a recent issue.

Articles are published in the order of their receipt, except in rare circumstances. Authors are supplied with 50 free offprints without covers. All of the cost of an author's alterations in galley proof is charged to the author. Priority in publication is given to articles whose authors assume the full cost of publication. In 1961 the publication cost is $30 per printed page.

All **regular** articles, **supplementary reports**, and **replication reports** should be submitted to the Editor-elect: David A. Grant: Department of Psychology, University of Wisconsin; Madison 6, Wisconsin. **Early publication** articles should be sent to the Editor: Arthur W. Melton; Department of Psychology, University of Michigan; Ann Arbor, Michigan until *July 1, 1962*. Thereafter they should be sent to the Editor-elect. Manuscripts must be accompanied by an abstract of 100 to 120 words typed on a separate sheet of paper. The abstract should conform to the style of *Psychological Abstracts*. Detailed instructions for preparation of the abstracts may be obtained from the Editor or from the APA Central Office.

HELEN ORR
Acting Production Manager

RUTH P. JAGOLINZER
Editorial Assistant

Second-class postage paid at Lancaster, Pa.

Send address changes to 1333 Sixteenth St. N.W., Washington 6, D. C. Address changes must reach the Subscription Office by the 10th of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee second-class forwarding postage. Other claims for undelivered copies must be made within four months of publication.

© 1962 by the American Psychological Association, Inc.

Journal of
Experimental Psychology

VOL. 63, NO. 3 MARCH 1962

EFFECTS OF PATTERN OF REINFORCEMENT AND VERBAL INFORMATION ON ACQUISITION, EXTINCTION, AND SPONTANEOUS RECOVERY OF THE EYELID CR [1]

THOMAS F. HARTMAN AND DAVID A. GRANT

University of Wisconsin

This study was an investigation of the effects of certain patterns of reinforcement on the acquisition, extinction, and spontaneous recovery of the eyelid CR. Following the different patterns of acquisition, the effects of informing *S* that the UCS would not be delivered during extinction and spontaneous recovery were determined. One general purpose of the study was to obtain information as to the degree that verbal information (and presumably self-instruction) may affect extinction and spontaneous recovery in the eyelid CR. Four patterns of reinforcement were investigated: 100% reinforcement, 50% random reinforcement, double-alternation reinforcement, and double-alternation reinforcement with a trial counter and complete knowledge of the sequence of reinforced trials.

Grant, Riopelle, and Hake (1950) found marked inhibition of responding to double alternation of reinforcement when compared with 50% random reinforcement. Since instructions regarding the nature of the task have been found to affect the rate of conditioning with the conditioned eyelid response (e.g., McAllister & McAllister, 1958; Miller, 1940), the conditioned salivary response (Razran, 1935), and the conditioned GSR (Cook & Harris, 1937), it seems logical to assume that information concerning the patterning of reinforcement would be of importance. To test this hypothesis a patterned group that would have complete information about the reinforcement sequence (double alternation given reinforcement "rule" and trial counter) was compared with a patterned reinforcement group, double alternation, that received no such aid in recognizing the pattern of reinforcement.

Norris and Grant (1948) found that instructing *S*s to try to inhibit their eyelid CRs produced slower acquisition and more rapid extinction. Instructions have also been found to

[1] This research was supported in part by a grant from the National Science Foundation and in part by the Research Committee of the Graduate School of the University of Wisconsin with funds provided by the Wisconsin Alumni Research Foundation.

affect the extinction of other CRs such as the conditioned GSR (Cook & Harris, 1937) and the conditioned finger withdrawal (Lindley & Moyer, 1961), but the extinction of some responses do not show this effect (Joyce, 1956). As the methodology of the eyelid CR has been more thoroughly explored than those of the GSR, hand-withdrawal response and verbal CRs, study of verbal parameeters of this response would be desirable in any event. One methodological complication of the eyelid CR offers an unusual opportunity for investigation of the effects of instructions. Because *S*s in eyelid conditioning experiments can be separated into voluntary and nonvoluntary responders solely on the basis of response form (Hartman & Ross, 1961; Spence & Ross, 1959), possible differential effects of patterns of reinforcement and of verbal information on these two classes of *S*s should prove most interesting.

Method

Apparatus.—The apparatus was essentially that described by Hartman and Grant (1960). The CS consisted of an 0.8-mL. increase in brightness of a 10-cm. circular mild glass disk. Ambient room illumination was approximately 1.0 mL. CS duration was 700 msec.; the UCS, a corneal air puff, occurred 540 msec. following the onset of the CS and was of 200 msec. duration. The intensity of the air puff was regulated by a 150-mm. column of mercury.

Procedure.—A 4 × 2 × 2 factorial design was used. The four acquisition conditions were 100% reinforcement (100%), 50% random reinforcement (R), double-alternation reinforcement (DA), and double alternation with a counter under the CS window to count the trials (DAC). Before extinction and before spontaneous recovery half of the *S*s were told that they would not receive the air puff for the remainder of that session. Acquisition, extinction, and spontaneous recovery conditions were orthogonal to each other.

All *S*s received 60 acquisition trials, 30 extinction trials, and a 30-trial spontaneous recovery test approximately 24 hr. later. The intertrial interval varied between 20 and 40 sec. with a mean of 30 sec. The trials were programed by means of a Western Union tape transmitter. The "random" condition was completely random except the last acquisition trial was always reinforced and the first and last halves of the acquisition series contained an equal number of reinforcements. Five random sequences were used.

Before beginning the initial experimental session *S*s were given the "neutral" instructions used at the University of Wisconsin laboratory. Those *S*s in Group DAC were also told what the reinforcement pattern in the session would be, and were quizzed until they could tell whether any given trial would include an air puff. The counter was not operative for the other groups but was in the field of vision. In addition all *S*s were told that they would be interrupted three times during the experimental session; those groups given no information about extinction were interrupted twice "to see if they were comfortable" and once to tell them when the session was half over. The groups given information about extinction were told that the third time they were interrupted, it would be to inform them that for the remainder of that session, they would get no more air puffs. If questioned about this statement *E* said that the meaning would become evident as the session progressed. Interruptions following Trials 5, 45, and 60 were made to minimize the probability of distraction due to the final interruption. On the day *S*s returned for the spontaneous recovery test, appropriate groups were told either that everything would be the same as on the previous day, or that during that session there would be no air puffs.

Subjects.—The *S*s were 112 women and 48 men from the elementary psychology courses. Assignment of *S*s to conditions was random within a replication of the 16 experimental conditions until it became necessary to modify assignment to equalize the number of each sex per group. Five *S*s were discarded: 1 failed to appear for the spontaneous recovery test, 3 adapted to the UCS, and 1 gave a record impossible to score due to an extremely high random blink rate.

Results

All eyelid closures with latencies from 200 to 550 msec. were considered anticipatory responses. In

some analyses, specified below, *Ss* giving voluntary form responses were removed from the data, using the criterion suggested by Hartman and Ross (1961); i.e., *Ss* with time derivatives of eyelid responses to light greater than 35% of the mean derivative of the reflex to the UCS for more than half of their responses were classified as voluntary *Ss*.

Acquisition—Figure 1 shows the mean percentage of anticipatory responses plotted as a function of blocks of five trials for acquisition. In the very first block of trials, the groups no longer appear the same; Group DAC responds at a lower level than the other groups. Group 100% approaches its asymptote faster than the other groups, and the level of response in Group DAC remains very low. The curves of Group R and Group DA virtually overlap during all of acquisition.

The derivative criterion classified 51 *Ss* as voluntary: 16 in Group 100%, 14 in Group R, 13 in Group DA, and 8 in Group DAC. A chi square test failed to disclose a significant effect of acquisition series on number of voluntary responders. When these voluntary *Ss* were eliminated the same acquisition trends were evident, the only difference being that the overall level of responding was less in all groups. An

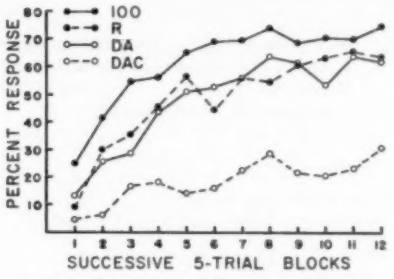

FIG. 1. Acquisition curves under four patterns of reinforcement. (Percentages of CRs are plotted against successive five-trial blocks for 100% reinforcement—100%, 50% random reinforcement—R, double alternation of reinforcement—DA, and double alternation with information and trial counter—DAC.)

analysis of variance of the total frequency of responses during the last half of the acquisition session gave highly significant *Fs* among reinforcement groups of 19.18 for the data of all *Ss* and 13.52 when voluntary *Ss* were eliminated, with error estimates based on 156 and 105 *df*, respectively.

Another group of analyses were performed on the terminal acquisition trials. Over the last 28 trials the reinforced and nonreinforced trials in Groups DA and DAC were analyzed. The percentage frequency of responses on reinforced and nonreinforced trials is presented separately for voluntary and nonvoluntary re-

TABLE 1

PERCENTAGE FREQUENCY OF RESPONSES GIVEN BY VOLUNTARY AND NONVOLUNTARY *Ss* ON REINFORCED AND ON NONREINFORCED TRIALS DURING THE LAST 28 ACQUISITION TRIALS

Response Type	Group DA			Group DAC		
	N	Reinforced Trials	Nonreinforced Trials	*N*	Reinforced Trials	Nonreinforced Trials
Voluntary	13	73%	67%	8	65%	32%
Nonvoluntary	27	51%	54%	32	22%	15%

TABLE 2

ANALYSES OF VARIANCE OF FREQUENCY OF RESPONSES IN 200–550 MSEC. RANGE ON REINFORCED AND NONREINFORCED TRIALS DURING LAST 28 ACQUISITION TRIALS

Group	Source	Voluntary *Ss*		Nonvoluntary *Ss*	
		df	*F*	*df*	*F*
DA	Reinforced-Nonreinforced	1	2.25	1	0.80
	Ss	12	19.36**	26	12.97**
	Error (*MS*)	12	(2.07)	26	(3.32)
DAC	Reinforced-Nonreinforced	1	11.32*	1	2.30
	Ss	7	1.02	31	4.10**
	Error (*MS*)	7	(7.56)	31	(3.59)

* $P = .05$.
** $P = .01$.

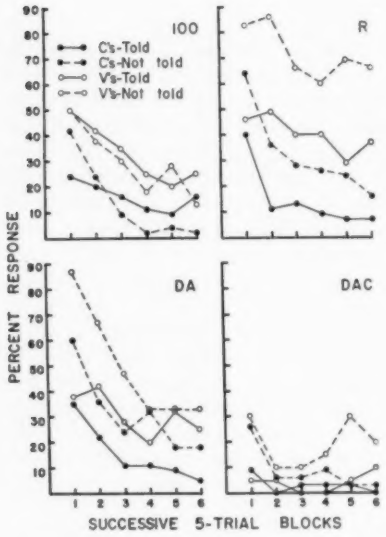

FIG. 2. Extinction of eyelid responses following four patterns of reinforcement during acquisition. (Percentages of CRs are plotted against successive five-trial extinction blocks. Each reinforcement group is subdivided into *Ss* giving CRs—Cs—and *Ss* giving voluntary responses—Vs—on the basis of the form of the response. These are further subdivided on the basis of whether they were informed—Told—when extinction procedures started or whether they were not so informed—Not Told.)

sponders in Groups DA and DAC in Table 1. There appear to be no systematic differences between responses to reinforced and nonreinforced trials in Group DA. In Group DAC, however, there is less responding on nonreinforced trials, and this is especially evident among the voluntary responders. The summaries of the analyses of variance are given in Table 2 where it is seen that the only significant difference between reinforced and nonreinforced trials was found among the voluntary responders of Group DAC. A notable absence of individual differences also occurred in this subgroup; every one of the 8 voluntary responders gave an equal or greater frequency of responses on reinforced trials than they did on nonreinforced trials. For these *Ss* the information about the reinforcement sequence and the aid of the counter definitely affected responsiveness.

Extinction.—Figure 2 presents the extinction curves for the nonvoluntary and voluntary *Ss* separately. The ordinate is percentage of responses in the 200- to 550-msec. interval; the abcissa is successive

five-trial blocks. Two major effects are evident in these curves: *Ss* told when extinction begins give fewer responses at the beginning of extinction than those *Ss* not given this information, and those *Ss* classified as voluntary responders on the basis of form give more responses in extinction than the nonvoluntary *Ss* regardless of the information given them.

Six separate analyses were made of the extinction data. Two analyses utilized the data of all *Ss*. The other four analyses considered either nonvoluntary *Ss* or voluntary *Ss* separately. With these last four analyses those *Ss* that did not make five anticipatory responses during the acquisition series were not included in the analysis. It was hoped that this procedure would reduce the error variance and partially compensate for the loss of power due to reduced numbers and unequal numbers per group in these analyses. The procedure used to adjust the bias caused by the unequal numbers of *Ss* per group was that outlined by Snedecor (1946, pp. 284–293).

Analyses of variance were performed on the last 30 trials within each acquisition reinforcement group to determine if those groups of *Ss* which subsequently received differential extinction and spontaneous recovery treatment were significantly different in acquisition. The largest F found was 2.09 ($df = 3/36$, $P > .05$). Separate analyses were performed for the data of all *Ss*, nonvoluntary and voluntary *Ss*.

TABLE 3

ANALYSES OF VARIANCE OF FREQUENCY OF RESPONSES IN 200–550 MSEC. RANGE FOR ALL *Ss* DURING EXTINCTION

Source	*df*	All Trials	First Five Trials Only
		F	*F*
Acquisition (A)	3	8.52**	10.87**
Extinction information (E)	1	3.28	8.31**
A × E	3	1.32	0.88
Error (*MS*)	152	(51.91)	(2.53)

** $P = .01$.

TABLE 4

ANALYSES OF VARIANCE OF THE EFFECT OF EXTINCTION PROCEDURE INFORMATION ON THE FREQUENCY OF RESPONSES IN 200–550 MSEC. RANGE FOR ALL *Ss* DURING EXTINCTION IN EACH REINFORCEMENT GROUP

Group	Source	*df*	All Trials	First Five Trials Only
			F	*F*
100%	Information	1	0.03	0.35
	Error (*MS*)	38	(63.21)	(2.87)
R	Information	1	4.66*	6.78*
	Error (*MS*)	38	(78.63)	(2.95)
DA	Information	1	0.18	1.13
	Error (*MS*)	38	(55.66)	(3.53)
DAC	Information	1	2.66	5.43*
	Error (*MS*)	38	(10.13)	(0.74)

* $P = .05$.

The analyses of variance of the extinction data of all *Ss* are presented in Table 3. When all trials are considered, the extinction information has no significant effect, but when the first five trials only are considered this information produces an effect significant at the .01 level.

Expectancy theory would predict that verbal information about onset of extinction would be most effective with Group R; therefore, it was decided to evaluate the effects of information separately in each acquisition group. These analyses are presented in Table 4. When the data for all trials are considered the effects of the extinction information are evident only with Group R, but

if only the first five trials are considered, both Group R and Group DAC show significant effects. In each group the effect of extinction information was to reduce responsiveness, especially on early extinction trials.

To determine if the differences between nonvoluntary *S*s and voluntary *S*s during extinction apparent in Fig. 2 were significant, separate analyses were performed for each response type. These analyses are presented in Table 5. If all trials were considered, extinction information had no significant effect for either group; when the first five trials only were considered, significantly fewer responses were found in the group told when extinction began for both the nonvoluntary and voluntary *S*s. Examination of each acquisition group separately produced only two significant *F*s, both with the voluntary *S*s in Group R. With the data on all trials, the *F* was 5.86, and with only the first five trials, the *F* was 5.98, both significant at the .05 level (error estimate based on 12 *df*). No significant *F* due to extinction information was found with Group DAC, but the test here with the voluntary *S*s was very insensitive as the error estimate was based on only 6 *df*.

TABLE 5

ANALYSES OF VARIANCE OF FREQUENCY OF RESPONSES IN 200–550 MSEC. RANGE DURING EXTINCTION FOR *S*s MEETING ACQUISITION CRITERION

Type of *S*	Source	*df*	All Trials	First Five Trials Only
			F	*F*
Non-voluntary	Acquisition (A)	3	2.15	4.33*
	Information on extinction (E)	1	2.43	8.71**
	A × E	3	0.81	0.08
	Error (*MS*)	66	(44.14)	(2.39)
Voluntary	Acquisition (A)	3	5.23**	3.66*
	Information on extinction (E)	1	3.67	6.70*
	A × E	3	1.18	1.40
	Error (*MS*)	41	(61.01)	(2.46)

* $P = .05$.
** $P = .01$.

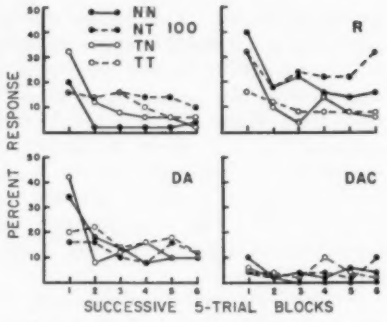

FIG. 3. Percentage frequency of eyelid responses plotted for successive 5-trial blocks during the spontaneous recovery test consisting of 30 extinction trials, following four patterns of reinforcement and a 30-trial extinction 24 hr. earlier. (After each acquisition procedure, *S*s are further subdivided into four groups according to whether they were informed—T—or not informed—N—of the absence of the UCS on initial extinction and on the spontaneous recovery test.)

Spontaneous recovery.—The spontaneous recovery data are plotted as Fig. 3; the axes are as in the previous figures, but the parameter separates groups according to whether they received extinction information (T) or not (N) on initial extinction and on the spontaneous recovery test. Comparing Fig. 3 with Fig. 2 there is evidently considerable spontaneous recovery, and again the effect of information at the beginning of the test session is evident. Those *S*s who were told there will be no air puffs at the beginning of the spontaneous recovery test session always give fewer responses than the corresponding group which receives no information. This is true regardless of whether *S*s had or had not been given extinction information on the occasion of first extinction.

Only two analyses were performed on the data gathered during spontaneous recovery; analyses of variance of all trials and the first five trials only for all *Ss*. Considering nonvoluntary and voluntary *Ss* separately was not feasible due to the small numbers which would be in some of the groups if the data were further divided. Table 6 contains summaries of the analyses of the spontaneous recovery data for all *Ss*, first for the data for all trials then for the first five trials only. As with the extinction data information has an effect only if the data for the first five trials are considered, not for the entire test session.

Discussion

Acquisition.—The outstanding result on the over-all acquisition curves is the lowered responding in Group DAC. In this respect, Group DAC resembles Group DA in an earlier study (Grant et al., 1950). Group DA of the present study gave far more responses than the DA group of the earlier one, and the only possible clue to the discrepancy lies in unsystematically recorded verbal reports of *Ss* in the two experiments. Whereas Grant, Riopelle, and Hake (1950) reported a large share of the *Ss* being able to report the double-alternation pattern, very few of the *Ss* in Group DA of this study could verbalize the double alternation in the reinforcement pattern. As pointed out below, perception of the pattern of reinforcement seems to be an important factor in reducing responsiveness.

TABLE 6

Analyses of Variance of Frequency of Responses in 200–550 Msec. Range for All *Ss* during Spontaneous Recovery

Source	*df*	All Trials	First Five Trials Only
		F	*F*
Acquisition (A)	3	4.20**	6.68**
Information on extinction (E)	1	0.53	0.02
Information on spontaneous recovery (SR)	1	0.20	7.47**
A × E	3	1.07	1.03
A × SR	3	0.17	0.62
E × SR	1	0.03	0.35
A × E × SR	3	0.23	0.12
Error (*MS*)	144	(29.96)	(1.77)

** $P = .01$.

In Table 1 were reported the percentages of responses in the CR latency range of *Ss* in Groups DA and DAC on reinforced and on nonreinforced trials. Each group was separated into the voluntary responders and the nonvoluntary responders on the basis of the objective criterion reported by Hartman and Ross (1961). Table 1 shows that neither the voluntary nor the nonvoluntary *Ss* responded differentially to the pattern reinforcement in Group DA. Furthermore, the general response rate was as high as in Group R. In Group DAC, however, there was differential responding to the reinforced and nonreinforced trials, and it followed a very interesting pattern. Only in the case of the voluntary *Ss* on the reinforced trials was the response level as high as that in Group DA. The nonvoluntary *Ss*, although they showed a slight but insignificant tendency to follow the reinforced-nonreinforced trial pattern, showed a marked general reduction in responsiveness to both the reinforced and nonreinforced trials. In fact, they were responding at less than 50% of the level attained by their fellow nonvoluntary *Ss* in Group DA. The voluntary *Ss* responded less than 50% as often on the nonreinforced trials as they did on the reinforced trials, and this result was statistically significant as shown in Table 2. Evidently, only the voluntary *Ss* were able to make effective use of the information and cues regarding the double alternation reinforcement patterns. But the tremendous reduction in CRs among the nonvoluntary *Ss* in

Group DAC was itself a very important finding. It appears to be a phenomenon over and above that of the general reduction in responsiveness obtained when nonreinforced trials are given as in Groups R or DA. Information as to whether a trial will or will not be reinforced, instead of assisting *S* in responding appropriately, seems to result in a generalized inhibition. That this is more than general distraction of *S* by the trial counter, etc., seems indicated by the fact that the voluntary *S*s are able to respond at a high rate on the reinforced trials in Group DAC. The experiment provides no evidence as to whether the generalized inhibition characterizes the double alternation with a reinforcement cue, or whether it is due to the reinforcement cue alone or to some other factors in the experiment.

Extinction.—The general extinction curves shown in Fig. 2 are similar to those usually obtained when it is borne in mind that an acquisition criterion of five anticipatory responses during the training trials was required for *S* to be included in the analysis. The relatively high response level of Group DA during extinction as compared with Group 100% and Group R indicates that randomness of reinforcement alone was not a factor in the resistance to extinction, although only in the voluntary *S*s of Group R was resistance to extinction found to persist through the entire extinction series. Presumably, Group DA, receiving a patterned reinforcement sequence, should be able to discriminate the onset of extinction more readily than Group R. This evidently was not the case, indicating that the DA *S*s could not use the regularity of the reinforcement pattern to detect the onset of extinction just as they could not use the pattern in responding differentially to reinforced and nonreinforced trials during acquisition.

The effect of extinction information was most pronounced in Group R, and least pronounced in Group 100%. The statistical tests of Table 4 indicate that the extinction information was most effective on the first five trials of extinction except in Group R where a persisting effect of the information could be demonstrated through the entire extinction process. Extinction information seemed to be no more effective for the voluntary *S*s than it was for the nonvoluntary *S*s. Here is an indication that the voluntary *S*s were not especially able to utilize information given as to extinction and to adopt "correct strategy" in their responses during extinction. Instead, they tend to continue responding at a fairly high level throughout the extinction series. They made as much use of the extinction information but no more use of it than did the nonvoluntary *S*s.

Two somewhat contradictory aspects of the responding of the voluntary *S*s seem to be demonstrated. First, during acquisition in Group DAC, the voluntary *S*s were able to respond differentially to the reinforced and nonreinforced trials when given appropriate information. Secondly, during extinction, the appropriate information as to the extinction procedure produced no greater effect on the voluntary *S*s than it did on the nonvoluntary *S*s. Instead, the voluntary *S*s continued to respond at a high level throughout extinction. In this latter aspect, the responsiveness of voluntary *S*s seemed to be characterized by high *D* (Spence, 1958). Unfortunately, because of the extremely rapid extinction in Group DAC, it is not possible to tell whether the same *S*s who responded differentially during reinforcement were able to utilize the extinction information. But it seems clear that under some circumstances, voluntary *S*s can utilize information to alter their responses according to a "correct strategy;" that in other instances, they are no more able to do so than other *S*s. Alternatively, voluntary responders may include some *S*s who tend to "respond appropriately" to stimulus patterns and some *S*s who merely respond at a high level.

Spontaneous recovery.—As usual (Grant, Hunter, & Patel, 1958; Hartman & Grant, 1960; Howat & Grant, 1958;

Porter, 1938, 1939; Prokasy, 1958), considerable spontaneous recovery was obtained on the spontaneous recovery test trials. These responses occurred primarily on the first extinction block and were followed by rapid extinction. The information as to whether reinforcement would be given during the spontaneous test trials was very important on the first five trials of the spontaneous recovery test regardless of whether *Ss* had been told or not been told about the extinction process in the first extinction procedure the day before. Evidently, the recovered CRs can still be affected by extinction information after an actual extinction process has been carried out. Furthermore, whether previous information on extinction had been given or not, no persisting effect of the information on first extinction was found during spontaneous recovery, as was shown by the curves in Fig. 3 and by the absence of significant *Fs* for extinction information and extinction by spontaneous recovery information interaction in Table 4. It appears that even after a first actual extinction followed by a 24-hr. recovery period, *Ss* are still able to respond differentially to the information that no reinforced trials will be given during the spontaneous recovery test. Something survives the first extinction that can be affected by verbal information about the extinction procedure.

Summary

The effects of 100% reinforcement, (Group 100%) 50% random reinforcement, (Group R) double-alternation reinforcement (Group DA), and double-alternation reinforcement with a trial counter (Group DAC) upon the acquisition, extinction, and spontaneous recovery of the conditioned eyelid response were studied. In addition, the effects of telling *S* at the beginning of extinction or spontaneous recovery that the air puff would no longer be presented were observed. All *Ss* received 60 acquisition trials, 30 extinction trials, and a 30-trial spontaneous recovery test 24 hr. later. The CS was a light and the UCS was a corneal air puff. The principal findings were as follows:

1. Groups DA and R produce the same degree of conditioning in acquisition, but Group DAC showed a considerably lower level of response than the other two groups reinforced 50% of the time.

2. In Group DA, neither the nonvoluntary nor voluntary *Ss* responded significantly more to the reinforced than to the nonreinforced trials; however, in Group DAC, the voluntary *Ss* responded significantly more to the reinforced trials.

3. Knowledge of cessation of the UCS produced a significant decrement of response during the first five extinction trials with both nonvoluntary and voluntary *Ss*. The voluntary *Ss* did not show immediate extinction when told the air puff would be omitted. Voluntary *Ss*, therefore, showed two contradictory aspects in their responding. They were able to respond significantly more to the reinforced trials in Group DAC, but they were not able to utilize knowledge of the cessation of the UCS during extinction to respond appropriately. Instead, they responded at a generally higher rate than the nonvoluntary *Ss* during extinction.

4. Considerable spontaneous recovery was found on the first five-trial block followed by rapid extinction. Knowledge that the UCS would not be given on the spontaneous recovery test reduced the number of responses on the test regardless of the occurrence of an earlier actual extinction series with or without verbal instructions as to the cessation of the UCS on the earlier extinction trials.

References

Cook, S. W., & Harris, R. E. The verbal conditioning of the galvanic skin reflex. *J. exp. Psychol.*, 1937, **21**, 202–210.

Grant, D. A., Hunter, H. G., & Patel, A. S. Spontaneous recovery of the conditioned eyelid response. *J. gen. Psychol.*, 1958, **59**, 135–141.

Grant, D. A., Riopelle, A. J., & Hake, H. W. Resistance to extinction and the pattern of reinforcement. I. Alternation of reinforcement and the conditioned eyelid response. *J. exp. Psychol.*, 1950, **40**, 53–60.

Hartman, T. F., & Grant, D. A. Effect of intermittent reinforcement on acquisition, extinction, and spontaneous recovery of the conditioned eyelid response. *J. exp. Psychol.*, 1960, **60**, 89–96.

Hartman, T. F., & Ross, L. E. An alternative criterion for the elimination of "volun-

tary" responses in eyelid conditioning. *J. exp. Psychol.*, 1961, **61**, 334–338.

HOWAT, M. G., & GRANT, D. A. Influence of intertrial interval during extinction on spontaneous recovery of conditioned eyelid responses. *J. exp. Psychol.*, 1958, **56**, 11–15.

JOYCE, J. N. The role of language as a cue in learning and extinction. *Dissertation Abstr.*, 1956.

LINDLEY, R. H., & MOYER, K. E. Effects of instructions on the extinction of a conditioned finger-withdrawal response. *J. exp. Psychol.*, 1961, **61**, 82–88.

MCALLISTER, W. R., & MCALLISTER, D. E. Effect of knowledge of conditioning upon eyelid conditioning. *J. exp. Psychol.*, 1958, **55**, 579–583.

MILLER, J. The effect of facilitatory and inhibitory attitudes on eyelid conditioning In E. R. Hilgard & D. G. Marquis (Eds.), *Conditioning and learning*. New York: Appleton-Century, 1940.

NORRIS, E. B., & GRANT, D. A. Eyelid conditioning as affected by verbally induced inhibitory set and counter reinforcement. *Amer. J. Psychol.*, 1948, **61**, 37–49.

PORTER, J. M., JR. The modification of conditioned eyelid responses by successive series of nonreinforced elicitations. *J. gen. Psychol.*, 1938, **19**, 307–323.

PORTER, J. M., JR. Experimental extinction as a function of the interval between successive nonreinforced elicitations. *J. gen. Psychol.*, 1939, **20**, 109–134.

PROKASY, W. F. Extinction and spontaneous recovery of conditioned eyelid responses as a function of amount of acquisition and extinction training. *J. exp. Psychol.*, 1958, **56**, 319–324.

RAZRAN, G. H. S. Conditioned responses: An experimental study and a theoretical analysis. *Arch. Psychol., N. Y.*, 1935, **28**, No. 191.

SNEDECOR, G. W. *Statistical method.* (4th ed.) Ames, Iowa: Iowa State Coll. Press, 1946.

SPENCE, K. W. A theory of emotionally based drive (*D*) and its relation to performance in simple learning situations. *Amer. Psychologist*, 1958, **13**, 131–141.

SPENCE, K. W., & ROSS, L. E. A methodological study of the form and latency of eyelid responses in conditioning. *J. exp. Psychol.*, 1959, **58**, 376–381.

(Received December 27, 1960)

Journal of Experimental Psychology
1962, Vol. 63, No. 3, 227-234

PATTERN RECOGNITION OVER DISTORTIONS, BY HUMAN SUBJECTS AND BY A COMPUTER SIMULATION OF A MODEL FOR HUMAN FORM PERCEPTION

LEONARD UHR
University of Michigan
CHARLES VOSSLER
System Development Corporation
AND JAMES ULEMAN
Harvard University

The recognition of single randomly generated, "meaningless" patterns has been studied during the past few years by a number of investigators (Attneave, 1957; Crook, Gray, Hanson, & Weisz, 1959; Vanderplas & Garvin, 1959b). Several experiments have further examined the abilities of human *S*s in recognizing such patterns over certain simple variations, such as random noise (Crook et al., 1959; Hillix, 1960), contour noise (Fitts & Leonard, 1957), and systematically continued transformations (LaBerge & Lawrence, 1957). But no experiments have been reported in which *S*s were asked to learn sets of variants of a pattern through experience with individual examples of these sets. Yet this is the typical procedure for computer "pattern recognition" simulation programs (Selfridge & Neisser, 1960; Uhr, 1962). It also presents a good experimental paradigm for the learning of patterns by human beings. Given N patterns and n variants of each pattern, *S* is presented with each of the nN particular pattern instances, asked to give it the correct name of the N names, and then told the correct name. This, in more systematic form, is what the child experiences during the natural experiments of his day-by-day learning. That is, the child is given particular instances of such things as "chair," "dog," the letter "A," or his mother's face; and he learns to build up general concepts, or pattern classes.

This sort of experiment, then, should serve as a convenient method for studying concept formation and perceptual learning of patterns in human *S*s. It should also give a convenient framework within which to examine the predictive fit of pattern recognition simulations by computer programs, when they purport to be embodiments of theoretical models of human form perception. The use of meaningless patterns for stimulus sets reduces the wealth of information that the human can bring to bear. It also allows for control and systematic variation of the parameters of the stimulus sets.

The main experiment examined the learning of pattern sets as a function of (*a*) interrelations between patterns, (*b*) practice, and (*c*) individual differences between *S*s (including the simulated model as an *S*). Subsidiary experiments examined the effects of different (*a*) durations of stimulus presentation, (*b*) size of pattern set, and (*c*) complexity of pattern.

In addition, three form perception

experiments were replicated using the computer as *S*. These were those of Attneave (1957) and Vanderplas and Garvin (1959b), who examined *S*s' abilities to make fine discriminations between individual patterns, and Fitts and Leonard (1957), who examined *S*s' abilities to generalize from individual patterns to variants.

The computer program is the first version of a model for form perception and concept formation. The program does not appear to take advantage of such trivially superior abilities of a digital computer as perfect rote memory or extreme arithmetic precision. Nor does it have built-in imprecision to simulate so-called imprecisions in human *S*s, such as memory decay. Finally, it should be noted that this is not a trivial test for computer programs of this sort, at the present stage of their development. In fact, relatively few programs are capable of even attempting to handle arbitrary new pattern arrays, and it is not known whether any would do appreciably better than the present program (Highleyman, 1961; Uhr, 1962).

1a. The "easy" set (taken from Crook et al., 1959).

1b. The "hard" set (taken from LaBerge & Lawrence, 1957).

FIG. 1. Stimuli used to make up the two experimental sets of five variants of each of five different patterns.

METHOD

Main Experiment

Subjects and design.—Human *S*s were tested in groups of 6 to 10, approximately half men and half women. For the analyses of variance, 3 men and 3 women were chosen randomly from each group. The *S*s were all paid volunteer college students. Each *S* was tested on two stimulus sets; half were given the "unrelated" stimulus set first, the other half were given the "interrelated" stimulus set first. A test on a stimulus set consisted of five learning trials through the entire set of 25 stimuli, i.e., five variants of each of the five patterns. The order of presentation was determined by using a random number table, with ordering different for each of the five trials.

Apparatus and stimuli.—Two sets of patterns were chosen that would be sufficiently difficult so that *S*s could not learn them immediately, but sufficiently learnable so that reasonably interesting learning curves would be generated. These stimulus sets are reproduced in Fig. 1.[1] Each set consisted of five pattern types, each pattern type with five variants. One set was chosen from Crook's eight-sided patterns (Fig. 1a). This was the "easy," "unrelated" set. The second set was chosen from LaBerge and Lawrence's

[1] We would like to thank M. N. Crook, of Tufts University, and D. H. Lawrence, of Stanford University, for their kindness in supplying us with usable copies of their original stimuli.

(1957) systematically varied patterns, in which the five different pattern types are actually related to one another (Fig. 1b). This was the interrelated or hard set. The "distance" as measured in an objective manner among members of the *same* pattern type (and more or less confirmed by the subjective judgments collected by LaBerge and Lawrence) is in certain cases *greater than* the distance *between* members of *different* pattern types.

Stimuli were projected, one at a time, by means of a reflector projector, onto a screen approximately 5 to 8 ft. in front of *Ss*. Stimulus size was approximately 3 in. square, which subtended roughly a .5° to 1° arc at the retina. These distances were chosen to to give optimal viewing conditions and still be comparable to the computer's resolving power. Pretesting was used to determine the most favorable lighting and viewing conditions, ones that would result in maximal learning for human *Ss*. Each stimulus was presented for 5 sec. Ten to 15 sec. after its removal, *Ss* were told what the correct name or type was. The following stimulus appeared 3 to 7 sec. after that.

Several subsidiary experiments were run as checks on some of these control conditions. (*a*) One subsidiary control group was given the same stimulus set under the same conditions, except for a 2-sec. presentation period. (*b*) An additional group of *Ss* was given a set of seven patterns, including the five that had been used for the "easy" set, in order to see to what extent short-term memory for too large a number of different pattern sets might be serving as a limiting factor for learning. (*c*) The effects of different complexities of stimuli were examined with a third group of *Ss*, who were given four-sided patterns. (Complexity was defined by the number of sides to the randomly generated stimulus.) In each of these subsidiary experiments, all other conditions were held constant. Results are presented as comparisons between the two pertinent groups involved.

Finally, a computer program designed to recognize patterns across their variants in a manner suggested by psychological and physiological data (Uhr & Vossler, 1961, 1962) was also tested on these same input patterns. Inputs were presented to the computer in the same order that they had been presented to *Ss*. The resolution of the inputs was also the same. Although there was no possibility of making a meaningful equation between computer program and *Ss* for such matters as duration of presentation of stimuli, illumination, and other viewing conditions, an effort was made to give *Ss* the best possible viewing conditions.

Briefly, the pattern recognition program works as follows: Unknown patterns are presented to the computer in discrete form, as a 20 × 20 matrix of zeros and ones. The program randomly generates measuring operators that simulate local neuron nets, by one of several random methods, and uses this set of operators to transform the unknown input matrix into a list of four 3-bit characteristics. These characteristics are then compared with lists of characteristics in memory, one for each type of pattern previously processed. As a result of similarity tests, the name of the list most similar to the list of characteristics just computed is chosen as the name of the input pattern. The characteristics are then examined by the program and depending on whether they individually contributed to success or failure in identifying the input, amplifiers for each of these characteristics are then turned up or down. This adjustment of amplifiers leads eventually to discarding operators which produce poor characteristics, as indicated by low amplifier settings, and to their replacement by newly generated operators.

Replications

All replication experiments used only the computer as *S*. Here again an attempt was made to equate computer and human conditions wherever possible. Attneave (1957) asked *Ss* to discriminate among members of a set of eight randomly produced variants of a prototype. The computer was tested on one of the several sets which Attneave found essentially equivalent.[2] Vanderplas and Garvin (1959b) asked *Ss* to learn a set of eight low-association-value patterns in a paired-associate procedure, in which the other member of the association was a nonsense syllable. Fitts and Leonard (1957) had *Ss* learn 12 different prototype patterns during three 6 to 10 min. self-study periods.[3] The 70% who passed the criterion of three out of four correct test trials were then tested

[2] The authors are grateful to F. Attneave for supplying us with his original stimuli.

[3] Fitts and Leonard (1957, p. 30) have presented results of their experiment in this mode, and P. Fitts has kindly supplied the authors with the original stimuli, along with details (Leonard & Fitts, unpublished manuscript) of their experiment.

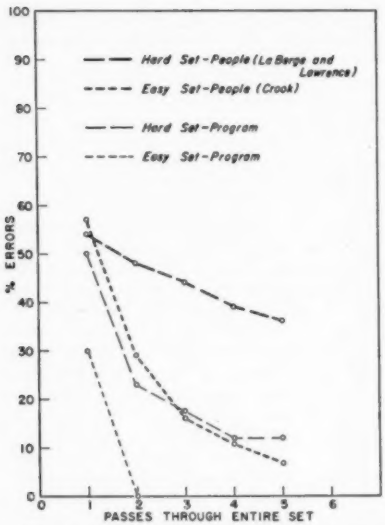

FIG. 2. Mean percentage errors for human *Ss* (from 6 to 10 *Ss* per point) compared with percentage errors for computer simulation program, on the easy and on the hard sets of stimuli (first presentation).

on generalization to new variants of the prototypes.

RESULTS

Main Experiment

Figure 2 presents mean errors on successive trials ("passes") for *Ss* and for the computer simulation program. Table 1 presents results from the analysis of variance on these data. The two pattern sets, and their two presentations, have been analyzed as a single dimension of four "conditions," because of limitations in available computer programs. The program's performance can probably most meaningfully be compared with the human performance in learning the *first* of the two sets learned, since the program does not accumulate broad general abilities and facility for new problems through experience in any way comparable to the human being. Therefore, it did not appear worthwhile to run a "second set" experiment using the program as *S*. Instead, Set 1 results were used as best estimates of what performance would have been on Set 2. (This would appear to be a conservative estimate, since it is saying in effect that no learning will occur, although, in fact, the program *is* capable of self-modification, and actually "learns" to recognize patterns originally through a process of this sort.) Differences between "conditions," between "passes," and between *Ss* were all significant beyond the .001 level.

TABLE 1

ANALYSIS OF VARIANCE OF ERRORS AS A FUNCTION OF PATTERN SETS, TRIALS, AND *Ss*

Source	Simulation Program as Contrasted with Mean Human Performance			Simulation Program as One of 7 *Ss*, along with Individual Human *Ss*		
	df	*MS*	*F*	*df*	*MS*	*F*
Between *Ss*	1	1265.63	16.40**	6	91.75	7.34***
Trials (T)	4	1163.21	15.08***	4	207.62	34.85***
Conditions (C)	3	1562.63	20.26***	3	539.05	43.16***
Ss × T	4	36.31	.47	24	5.69	.96
Ss × C	3	207.42	2.69	18	12.49	2.10*
T × C	12	36.81	.48	12	17.54	2.95**
Ss × T × C	12	77.11		72	5.94	
Total	39			139		

Note.—The computer simulation program is included as an *S*.
* $P = .05$.
** $P = .005$.
*** $P = .001$.

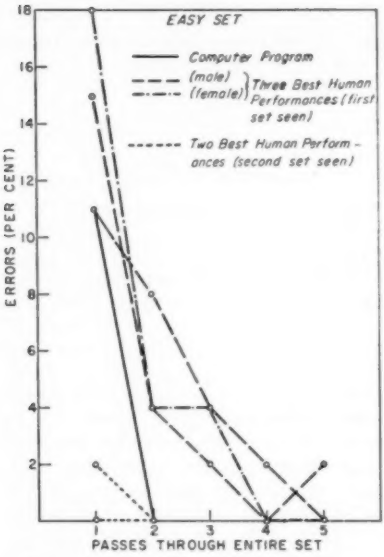

FIG. 3. Computer simulation program performance compared with best individual human performances on the easy set of stimuli.

In particular, the difference between the simulation program and the mean human performance was significant beyond the .005 level, the simulation's performance being consistently better.

Figure 3 presents results, for the unrelated pattern set, for the best of the individual *Ss*, and for the computer program. On this set of patterns, some of the human *Ss* when they had already had the experience provided by the other set of patterns, outperformed the simulation program. Figure 4 presents results, for the interrelated pattern set, for the best of the individual *Ss*, and for the program. On this more difficult set of patterns, the best human performances do not exceed the performance of the program. Nor, in fact, does the best human performance appear to improve very markedly as a result of the additional experience afforded *Ss* doing this task as their second task. That is, most of the improvement in the overall group performance appears to result from the improvement of the poorer members of the group.

Figures 5, 6, and 7 present results, for human *Ss* alone, for (*a*) different durations of presentation of stimuli—for 2-sec. and 5-sec. durations; (*b*) different size of pattern sets—for five and for seven patterns; and (*c*) different pattern complexities—four-sided and eight-sided patterns. There were no significant effects attributable to either number of pattern sets or duration of stimulus presentation (as varied in this experiment). Complexity of patterns did, however, tend to affect recognition, as indicated by the interaction effect (significant at the .05 level) between complexity and *Ss*.

Table 2 presents results of the

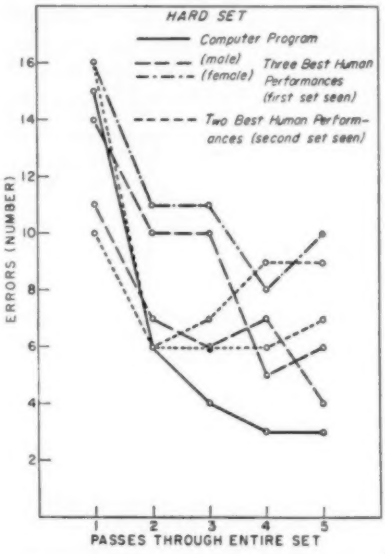

FIG. 4. Computer simulation program performance compared with best individual human performances on the hard set of stimuli.

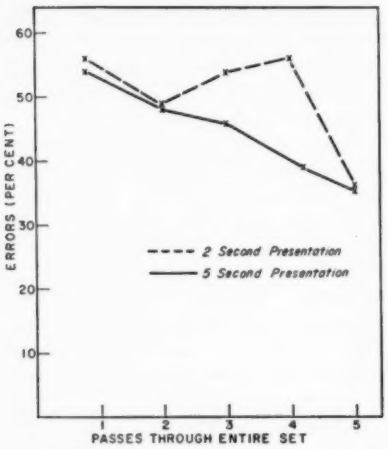

FIG. 5. Effects of duration of stimulus presentation. (Dotted line presents results—10 *Ss*—in learning when patterns were presented for a 2-sec. period. Solid line presents results—9 *Ss*—when patterns were presented for a 5-sec. period, all other conditions comparable.)

replications. For a more accurate comparison of Vanderplas and Garvin's (1959a, 1959b) results with ours, the error due to *Ss*' learning nonsense syllables would have to be excluded since the computer program dealt only with the visual patterns. But, since the program made no errors at all on its second pass through the set, we have not bothered to make such an adjustment. On the Fitts and Leonard (1957) set, the program learned the prototypes to a 75% level after its first experience, and to a 100% level after its second. It gen-

TABLE 2

PATTERN RECOGNITION ABILITIES OF HUMAN *Ss* AND OF COMPUTER SIMULATION OF A MODEL FOR HUMAN FORM PERCEPTION

Experiment	Mean Performance for Human *Ss*	Performance of Simulated Model
Attneave (1957) (cumulative errors to learning)	29.75	3.00
Vanderplas & Garvin (1959b) (2nd trial correct)	3%	100%
Fitts & Leonard (1957) (1st exposure percentage correct)	62%	100%

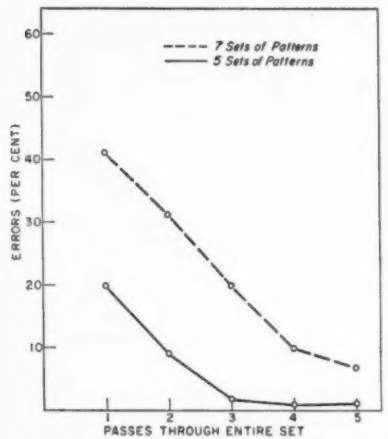

FIG. 6. Effects of size of pattern set. (Dotted line presents results—mean scores for 10 *Ss*—in learning seven sets of patterns. Solid line presents results—9 *Ss*—for five sets of patterns, all other conditions comparable.)

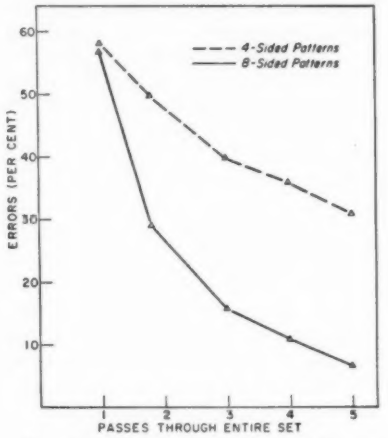

FIG. 7. Effects of number of sides to a pattern. (Dotted line presents results—8 *Ss*—in learning four-sided patterns. Solid line presents results—6 *Ss*—for eight-sided patterns, all other conditions comparable.)

eralized with 76% accuracy after the first experience, and with 100% accuracy after the second, as contrasted with human *S*s, who generalized with 62% accuracy after 1 hr. of learning.

Discussion

These results suggest that the simulation program may well be performing at a higher level as a "pattern recognizer" than can be achieved by the human *S* when he is divested of some of his advantages based upon past experience and contextual information. This, of course, is a very narrow test of the program's abilities, nor is it known to what population of patterns these results might be generalized. However, the program was not designed with these pattern sets in mind, or even with this general type of pattern set in mind. At this stage of experimentation, it would seem reasonable to simply continue to test such programs out on as wide a range of stimuli as possible.

The fact that the program performs better than human *S*s in this experiment should not be taken very seriously. The simulation is capable of coping with only a part of the perceptual problem. It would not be at all unlikely, or uninteresting, to find that, as the program's generality increases, its abilities at any particular task decreases as a direct and intrinsic function of its greater powers in other directions.

Further experiments of this sort should serve at least two purposes. First, by varying such things as the number of patterns, the number of variants of each pattern, the distance between these variants, the way the individual stimuli sample the pattern's total population, and the distance between patterns, a rather rich domain of behavior closely related to what we have traditionally named "perceptual learning" and "concept formation" should be thrown open to assessment. By testing performance on new variants, different from those originally learned, generalization of classes and concepts can be studied. Second, by comparing the simulation program's performance with human performance as a function of such variations, a beginning can be made in attempting to fit some of the details of the curves generated by the simulated model to the curves generated by human *S*s.

Summary

Human *S*s, and a computer simulation program of a model for form perception, were examined for their behavior in learning to respond with the proper name for a pattern type over variant examples of the pattern. In the main experiment, five pattern types of five variants each were presented to all *S*s (including the simulation program). Four "conditions" were examined: (*a*) unrelated patterns, (*b*) interrelated patterns; each presented as the first, and as the second, set. Differences between *S*s, and between "passes" through the set (learning trials) were also examined. All of the above variables showed differences significant beyond the .001 level. In particular, the simulation of the model performed at a significantly lower error rate (beyond the .005 level) than did the group of human *S*s.

The effects, on human *S*s, or varying (*a*) duration of stimulus presentation, (*b*) size of the set of patterns, and (*c*) complexity of patterns were also examined. Only complexity of pattern affected performance significantly.

In addition, three experiments by other investigators who tested human *S*s were replicated, to test the computer model. The model outperformed human *S*s in all cases.

References

Attneave, F. Transfer of experience with a class-schema to identification-learning of patterns and shapes. *J. exp. Psychol.*, 1957, **54**, 81–88.

Crook, M. N., Gray, F. E., Hanson, J. A., & Weisz, A. The effect of noise on the perception of forms in electro-visual display systems: A set of irregular forms. Tufts University, Inst. Appl. Exper. Psychol., 1959.

Fitts, P. M., & Leonard, A. J. Stimulus correlates of visual pattern recognition: A probability approach. Laboratory of Aviation Psychology, Ohio State University, 1957.

HIGHLEYMAN, W. H. A note on optimum pattern recognition systems. *IRE Trans. Electron. Comp.*, 1961, No. EC-10, 287–288.

HILLIX, W. A. Visual pattern identification as a function of fill and distortion. *J. exp. Psychol.*, 1960, **59,** 192–197.

LABERGE, D. L., & LAWRENCE, D. H. Two methods for generating matrices of forms of graded similarity. *J. Psychol.*, 1957, **43,** 77–100.

SELFRIDGE, O. G., & NEISSER, U. Pattern recognition by machines and men. *Scient. American*, 1960, **203,** 60–68.

UHR, L. "Pattern recognition" computers as models for form perception. *Psychol. Bull.*, 1962, in press.

UHR, L., & VOSSLER, C. A pattern recognition program that generates, evaluates and adjusts its own operators. *Proc. Western Joint Computer Conf.* 1961, **19,** 555–561.

UHR, L., & VOSSLER, C. A computer simulation of pattern perception and concept formation. In, *Proceedings of the bionics symposium: 1961.* Ithaca, N. Y.: 1962, in press.

VANDERPLAS, J. M., & GARVIN, E. A. The association value of random shapes. *J. exp. Psychol.*, 1959, **57,** 147–154. (a)

VANDERPLAS, J. M., & GARVIN, E. A. Complexity, association value, and practice as factors in shape recognition following paired-associate training. *J. exp. Psychol.*, 1959, **57,** 155–163. (b)

(Early publication received June 16, 1961)

Journal of Experimental Psychology
1962, Vol. 63, No. 3, 235–243

ON THE HETEROGENEITY OF STIMULUS AND RESPONSE ELEMENTS IN THE PROCESSING OF INFORMATION [1]

WILLIAM C. HOWELL

Ohio State University

Although it has been assumed that ". . . one or both of two simultaneous performances will show some impairment" (Woodworth & Schlosberg, 1955, p. 88), exceptions to this generalization appear to be many and varied. Klemmer (1956, 1958), for example, has noted that such an impairment is contingent upon task complexity or difficulty. Other investigators have pointed out that performance efficiency in multiple activities is a function of factors such as sense channel differences (Mowbray, 1953), task repetitiveness (Bahrick, Noble, & Fitts, 1954), input redundancy (Broadbent, 1958), and combined task training (Bahrick et al., 1954; Briggs & Waters, 1958). Finally, industrial research has indicated that a certain amount of alternation among tasks, particularly widely differing tasks, is actually beneficial to worker output (Ghiselli & Brown, 1948, pp. 420–423).

The number of variables contributing to the divergence of results in multiple-task studies is probably quite large. None is more basic, however, than the failure of investigators to distinguish between two very different multiple-task situations. One is the *time-sharing* task arrangement in which multiple stimuli, each with its associated response, occur at the same point in time (Bahrick et al., 1954; Briggs & Howell, 1959; Mowbray, 1953). The other involves only the alternation, in some fashion, of stimuli or S-R units defining the tasks and hence may be termed *task heterogeneity* (Baker, Wylie, & Gagné, 1951; Klemmer, 1956, 1958). It is not inconceivable that such dissimilar situations would lead to conflicting results. Temporal coincidence (i.e., time sharing) would virtually assure some decrement in one of the tasks, whereas alternation (i.e., task heterogeneity) could occur without impairment to either.

The present research was concerned with the influence of one of these variables, task heterogeneity, upon simple information-processing performance. Heterogeneity was provided by two stimulus and two response classes, the former distinguished on the basis of sense modality (audition vs. vision) and the latter on the basis of muscle group (verbal vs. manual response mode). Combining these elements in various ways, it was possible to obtain the following types of heterogeneity: (*a*) homogeneity (or zero heterogeneity) in which one stimulus class is paired with only one response class; (*b*) stimulus heterogeneity in which the two stimulus classes are paired with a single response class; (*c*) response heterogeneity in which the two response classes are paired with a single stimulus class; and (*d*) task hetero-

[1] This research was carried out in the Laboratory of Aviation Psychology and was supported by the United States Air Force under Contract No. AF 33(616)-3612, monitored by the Aerospace Medical Division. Permission is granted for reproduction, translation, publication, use, and disposal in whole or in part by or for the United States Government.

geneity in which both stimulus classes are used, each paired with its own response class. The plan for evaluating heterogeneity effects involved comparisons of the efficacy of specific S-R pairs to convey information when included in a heterogeneous task with their efficacy in this regard when included in a homogeneous task. Since task heterogeneity may be considered a combination of heterogeneous stimulus and response elements, a further objective was an evaluation of the relative contribution of stimulus and response heterogeneity to the over-all task heterogeneity effect.

A distinguishing feature of the current approach is that the method for producing heterogeneity (of stimuli, of responses, or of the task) created no change in the amount of information presented. In the typical time-sharing or heterogeneity investigation, the amount of stimulus information increases as tasks are added, so that performance with respect to the individual components may decrease while over-all information transmitted remains constant or even increases. In the present design, however, any changes in information transmitted must be attributed to heterogeneity as defined.

Experiment I

The first experiment was to establish base performance values for each of four tasks: (*a*) visual stimulus with manual response, (*b*) auditory stimulus with manual response, (*c*) visual stimulus with verbal response, and (*d*) auditory stimulus with verbal response. It should be recognized that such base values, obtained under homogeneous task conditions, were necessary both as a means of evaluating S-R compatibility effects within specific subtasks and as a basis of comparison for the heterogeneous tasks undertaken in the subsequent experiments.

Method

Apparatus and stimuli.—The OSU serial discrimeter was employed to present the stimuli and to score responses. A description of this device has been given elsewhere (Alluisi & Muller, 1958). Functionally, the apparatus permitted the preprograming of six stimulus events in any order desired for a sequence of 100 presentations. The *S* responded to each of these stimuli by pressing one of six response keys (manual response condition) or by pronouncing one of six numbers to activate a voice key (verbal response condition). The activation of either a manual or a verbal response key initiated the presentation of the succeeding stimulus in the sequence with the result that *S* maintained complete control of stimulus pacing. In addition, the activation of each manual key was recorded on one of 36 counters arranged in a 6×6 stimulus-response matrix. This scoring arrangement enabled *E* to record directly total errors, as well as individual response frequencies for use in computing entropy values. For verbal response trials it was necessary for *E* to record each response manually.

With the presentation of the first stimulus in each trial, a 1/1000-min. timer was activated which continued to run until the last response in the sequence had been completed. Thus, it was possible to record the over-all time required for the completion of all 100 responses in a sequence and from this to compute the average response time per stimulus presentation.

The six visual and six auditory stimuli used in the study are described fully in Table 1. The visual stimuli were selected from a list of colors demonstrated by Halsey and Chapanis (1951) to be easily discriminable. The stimuli were projected from the rear onto the center of a 10-in. diameter opal glass screen by means of a multiple-unit optical projector. The viewing distance for *S* was approximately 28 in., with the diameter of each color target subtending 31′ of visual angle at the eye.

The six tones used as auditory signals were selected on the basis of psychophysical judgments such that (*a*) the middle frequency bisected the distance in pitch between the two apparatus-imposed extremes, (*b*) the high loudness level was well below the discomfort threshold, (*c*) the low loudness level

TABLE 1

VISUAL AND AUDITORY STIMULI AND ACCOMPANYING RESPONSES: EXP. I

Stimuli				Responses	
Visual		Auditory		Verbal	Manual[a]
Color Name	Wratten Filter	Frequency (cps)	Intensity (db.)		
Light Red	24	100	60	1	LIII
Dark Blue	45a	500	53	2	LII
Light Green	52	6000	58	3	LI
Orange	106	100	90	4	RI
Violet	34a	500	85	5	RII
Gray	ND/0.6	6000	93	6	RIII

[a] Designations refer to the digit numbers (e.g., LIII means third digit on the left hand) and their associated keys.

was well above absolute threshold, and (*d*) loudness levels were matched at all frequencies. All tones were produced by individually preset electronic signal generators (Eico Model 377), the outputs of which were fed into a set of Telephonics headphones worn by *S*.

Subjects.—A total of 36 male students served in this experiment, each of whom was assigned at random to one of four groups. Each group, therefore, contained 9 *S*s, all of whom were trained to perform under one of the four homogeneous S-R conditions: visual-verbal (V-V), visual-manual (V-M), auditory-verbal (A-V), and auditory-manual (A-M).

Procedure.—In this, as well as in the succeeding experiments, all *S*s received the benefit of 6 practice trials (600 S-R pairings) prior to the performance on 3 experimental trials (300 S-R pairings). The amount of training provided was sufficient to bring all *S*s to a comparable level of performance. That this level was probably short of asymptotic is suggested by findings which demonstrate that improvement on such tasks continues for at least 50 100-stimulus trials (Alluisi & Martin, 1957). In both the training and experimental trials *S* was instructed to respond as rapidly and accurately as possible. Each training trial was followed by a demonstration of the S-R pairings to prevent *S* from establishing incorrect associations. The training trials were administered in blocks of three with the same random sequence being repeated throughout each block. Experimental trials, on the other hand, were administered singly with each sequence drawn at random from a population of 27 such sequences. Rest periods of 3-min. duration were interposed between experimental trials and between training blocks; periods of 1 min. were inserted between training trials within blocks. All training and experimental trials were completed in a single session, the duration of which was approximately 1 hr.

Results

Performance under each of the four conditions was evaluated using two information measures: information transmitted (H_t) and information transmitted per second (H_t/sec). Since errors were few and well distributed, the former measure consti-

TABLE 2

MEAN INFORMATION TRANSMITTED (H_t) PER RESPONSE AND RATE OF INFORMATION TRANSMITTED (H_t/SEC) SCORES OBTAINED UNDER THE FOUR HOMOGENEOUS CONDITIONS OF EXP. I

S-R Pairing	H_t (Bits)[a]		H_t/Sec	
	Mean	*SD*	Mean	*SD*
A-V	2.50	.05	1.68	.28
V-V	2.49	.04	1.94	.18
A-M	2.44	.11	2.06	.26
V-M	2.47	.01	2.48	.13

[a] Maximum possible H_t was 2.58 bits per response.

tutes a reasonable index of response accuracy, whereas the latter reflects both speed and accuracy of performance.

Considering first the H_t measure, it is apparent from Table 2 that only small differences occurred among S-R pairings with regard to accuracy. In fact, in no case did the differences exceed 2.4% of the absolute amount of information transmitted. An analysis of variance performed on these data confirmed the lack of statistical significance of the differences.

Turning to the H_t/sec measure, the results clearly indicate the superiority of the manual over the verbal response mode, and of the visual over the auditory stimulus mode. The statistical significance of the differences between means shown in Table 2 was demonstrated using a nonparametric technique developed by Alluisi (1956) for the purpose of testing analysis of variance hypotheses. Use of a nonparametric analysis was prompted by a failure of the H_t/sec data to meet the homogeneity assumption of the parametric analysis of variance. This analysis yielded χ^2 of 7.111 ($df = 1$, $P < .01$) for Stimulus Class, 7.111 ($df = 1$, $P < .01$) for Response Class, .445 ($df = 1$, *ns*) for S × R, and 14.667 ($df = 3$) for Total.

It is interesting to note that despite the significance of these major effects, the S × R interaction does not approach significance. Thus, it may be assumed that the S-R compatibility of the four tasks is comparable and hence can safely be ignored when these tasks are combined into various heterogeneity conditions as in the succeeding experiments. It must be added that the common level of S-R compatibility is not particularly high, for the transmission rates never exceeded 2.5 bits/sec.

Experiment II

Experiment II was concerned with the evaluation of task heterogeneity as it affects information-processing performance. Heterogeneity was introduced into the task scheme by the simple means of combining, as subtasks, equal portions of the homogeneous tasks described in Exp. I. Thus, two distinct heterogeneous tasks were available: visual-verbal, auditory-manual (VV-AM); and visual-manual, auditory-verbal (VM-AV).

The purpose of this experiment was to evaluate heterogeneity effects through comparison of performance on the heterogeneous tasks with performance on the component homogeneous tasks. The latter values, of course, were available from the results of Exp. I. Thus, it was intended that the average scores for each pair of subtasks obtained under homogeneous task conditions serve as the basis of comparison for the same subtasks performed under heterogeneous task conditions. Any differences obtained between these scores would then be attributable solely to heterogeneity effects.

Method

The apparatus and procedures used were the same as those described in Exp. I. Only two groups of *S*s were required for this experiment, one for each of the two task heterogeneity conditions (VV-AM and VM-AV). As in Exp. I, 9 male students were assigned at random to each group.

The specific S-R pairings used in the task heterogeneity conditions are summarized in Table 3. It may be noted that for both Conditions VV-AM and VM-AV the pairings were identical to those employed in Exp. I; that is, three S-R pairs were selected from each of the homogeneous tasks and combined to form two six-alternative heterogeneous tasks. In neither of these heterogeneous tasks was there any deviation from the stimulus sequencing arrangement used in Exp. I. Hence, auditory and visual stimuli

TABLE 3

HETEROGENEOUS STIMULUS-RESPONSE PAIRS USED IN EXP. II

Group I (VM-AV)		Group II (VV-AM)	
Stimuli	Responses	Stimuli	Responses
100 cps, 60 db.	1 (verbal)	Light red	1 (verbal)
500 cps, 53 db.	2 (verbal)	Dark blue	2 (verbal)
6000 cps, 58 db.	3 (verbal)	Light green	3 (verbal)
Orange	RI (manual)	100 cps, 90 db.	RI (manual)
Violet	RII (manual)	500 cps, 85 db.	RII (manual)
Gray	RIII (manual)	6000 cps, 93 db.	RIII (manual)

(calling for verbal and manual responses) were randomly interspersed within each trial sequence. In no case were S-R elements permitted to overlap in time.

Results

As may be seen in Table 4, the differences in H_t/sec obtained between the two heterogeneity conditions were small: the average rate of transmission values were 2.38 bits/sec for Cond. VV-AM and 2.26 bits/sec for Cond. VM-AV. The corresponding average values of H_t for the two conditions also were not dissimilar: 2.51 bits for VV-AM and 2.43 bits for VM-AV. Tests of these differences by the t statistic confirmed their failure to achieve significance.

In view of this lack of significance, the data from the two task heterogeneity conditions were averaged for comparison with mean homogeneous performance (obtained by averaging the scores in Table 2). The difference between mean H_t/sec scores for the homogeneous and heterogeneous task conditions was .29 bits/sec, and a Mann-Whitney U test applied to this difference achieved statistical significance ($P < .01$). Inspection of the accuracy data revealed that both heterogeneous and homogeneous conditions yielded identical mean H_t scores of 2.47 bits.

EXPERIMENT III

In view of the significant influence of task heterogeneity on rate of information transmission, it became meaningful to investigate the relative importance of stimulus and response factors in determining the over-all effect. The purpose of Exp. III, then, was to manipulate stimulus and response heterogeneity independently, comparing each to the basic conditions of Exp. I as well as to the task heterogeneity results of Exp. II.

Method

As in the task heterogeneity conditions (Exp. II), there were available two distinct tasks for testing both the stimulus and response heterogeneity effects. For stimulus heterogeneity these were auditory visual-verbal (AV-V) and auditory visual-manual (AV-M) S-R combinations; for the response heterogeneity conditions they were auditory-

TABLE 4

MEAN H_t AND H_t/SEC SCORES OBTAINED UNDER THE TWO HETEROGENEOUS TASK CONDITIONS

Cond.	H_t (Bits)		H_t/Sec	
	Mean	*SD*	Mean	*SD*
VV-AM	2.51	.08	2.38	.17
VM-AV	2.43	.04	2.26	.17
Mean	2.47	.06	2.32	.17

TABLE 5

MEAN H_t AND H_t/SEC SCORES OBTAINED UNDER STIMULUS AND RESPONSE HETEROGENEITY CONDITIONS

Task Components	H_t		H_t/Sec	
	Mean	*SD*	Mean	*SD*
Stimulus Heterogeneity				
AV-V	2.46	.07	1.90	.25
AV-M	2.48	.08	2.54	.15
Mean	2.47	.08	2.22	.20
Response Heterogeneity				
V-MV	2.49	.11	2.21	.31
A-MV	2.45	.10	2.12	.41
Mean	2.47	.10	2.16	.36

manual verbal (A-MV) and visual-manual verbal (V-MV) combinations.

The apparatus and procedures used were identical to those used in Exp. I and II. Again, 9 male students served as *Ss* for each condition. In Cond. AV-V and AV-M the stimuli used were 100, 500, and 6000 cps tones and orange, violet, and gray lights. The verbal responses to these stimuli were 1–6, respectively, and the manual responses LIII to RIII, respectively, as listed in Table 1. In Cond. A-MV and V-MV the Responses 1, 2, 3, RI, RII, and RIII were paired with all six auditory or all six visual stimuli as listed in Table 1.

Results

Mean H_t and H_t/sec scores for both stimulus and response heterogeneity conditions are given in Table 5. It is obvious that no reliable H_t differences occurred among task components, and, further, that these scores are no different from those obtained in Exp. I and II (all values approximate 2.47 bits). The difference between mean H_t/sec scores on the two stimulus heterogeneity conditions was large, the values being 1.90 bits/sec and 2.54 bits/sec for AV-V and AV-M, respectively. A t test applied to this difference indicated that it was statistically significant ($P < .01$). On the other hand, the difference between the two response heterogeneity means was not sufficiently large to approach statistical significance ($t = .53$). As a result, the two response heterogeneity conditions (V-MV and A-MV) were combined in their comparison with homogeneous conditions, whereas the two stimulus heterogeneity conditions (AV-V and AV-M) were considered independently.

The stimulus and response heterogeneity effects were again evaluated through comparison with homogeneous base conditions: mean scores on the heterogeneous tasks were compared with mean scores for the component subtasks performed under homogeneous conditions (Table 2). Of the two heterogeneous stimulus tasks only AV-M yielded a score appreciably above the homogeneous counterpart: the difference was .28 bits/sec, significant ($P < .01$) in terms of both t and U tests. Verbal performance, although insignificantly ($P > .10$) superior under the heterogeneous stimulus condition, was poor for both homogeneous and heterogeneous tasks (1.80 and 1.90 bits/sec, respectively). Turning to response heterogeneity, the mean of 2.16 bits/sec was only .13 bits/sec greater than that for the homogeneous base condition. Both t and U tests confirmed the insignificance of this difference ($P > .10$).

Comparisons were also made between average stimulus or response heterogeneity performance and that obtained for task heterogeneity in Exp. II. Although neither of the component heterogeneity conditions (i.e., stimulus or response) yielded H_t/sec values as high as those obtained under task heterogeneity, the differences were not statistically reliable.

Discussion

The results clearly indicate that task heterogeneity is beneficial to simple information-processing performance under conditions of low S-R compatibility and relatively little training. The findings with regard to the relative contribution of stimulus and response heterogeneity to the over-all task heterogeneity effect may also be stated simply: stimulus heterogeneity appears to increase information transmission significantly over the homogeneous control condition, whereas response heterogeneity leads to no reliable improvement. Whether the stimulus component is sufficient to account for the complete task heterogeneity effect is not clearly shown by the data. Statistically, there is no significant difference between the H_t/sec values for stimulus and task heterogeneity ($P > .05$), although the mean difference is fairly large in favor of task heterogeneity (.10 bits/sec). Further, the coding scheme employed in one of the stimulus heterogeneity conditions (AV-M) was such that it provided a useful cue for left- vs. right-hand responses. The relatively high transmission rates obtained under this condition may thus be somewhat spurious. It is probable, then, that the heterogeneity of stimuli is one factor contributing to the high rate of transmission associated with task heterogeneity. More conclusive, however, is the indication that stimulus heterogeneity is not the *only* contributing factor.

That such findings appear to be in distinct disagreement with many of the other studies encountered in this area may be attributed partially to differences in emphasis and design as suggested above. However, two investigations bear a sufficient degree of similarity to the present work to warrant further comparison. Klemmer (1956) reported that when *S* is forced to alternate regularly between visual and auditory signals more rapidly than once every 2 sec., his performance on a key pressing task declines sharply. In a later study, Klemmer (1958) was able to demonstrate the same effect for conditions of random alternation between visual and auditory channels.

To what, then, can be attributed the marked difference between the present results and those reported by Klemmer? The most obvious explanation appears to lie in the nature of the S-R pairs: whereas in the present research each response was associated with a single stimulus under any given condition, Klemmer used only three response elements for all six stimuli (three visual and three auditory). As a result, each response was associated with one visual and one auditory stimulus. It appears, then, as if a change in signal mode in some way combined briefly the tendencies for error on one channel with the comparable tendencies for error on the other channel without appreciably strengthening the association of either set of signals with the correct responses. In the present situation, of course, error tendencies would have little opportunity to combine since each signal was associated with its own response. Such an explanation, however, does not appear sufficiently strong to account completely for the large differences observed between the two studies.

Another explanation, also limited, is based upon the fact that Klemmer used forced-pacing of stimuli, whereas the present investigation permitted self-pacing by *S*. Of the two modes of presentation, self-pacing has been shown to yield a higher level of performance (Alluisi & Muller, 1958), although a full description of forced- vs. self-pacing differences has yet to be advanced. Considering the difference between the Klemmer results and those of the present investigations, then, it appears that the independence of S-R units and the manner in which stimuli are paced may determine the extent as well as the direction of the heterogeneity effect: for independent, self-paced S-R pairs task heterogeneity leads to improved performance whereas for nonindependent, forced-paced S-R pairs it incurs a decrement.

Assuming that the above explanation is sufficient to account for the differences reported, the problem of interpreting the beneficial heterogeneity effect obtained in the present research still remains. Why, to state the problem somewhat differently, should a simple task be more difficult to perform than a slightly more complex one?

Reasonable answers may be found within the traditional areas of learning and motivation. The learning interpretation relies upon the concept of stimulus generalization: since the introduction of stimulus or task heterogeneity divides the stimulus population into two distinct classes, the extent of stimulus generalization with respect to each signal should be reduced. As a result, it would be expected that the strength of incorrect or generalized associations (which associations control error scores as well as latency of responding) would likewise be reduced. The motivation hypothesis is less precisely put since it makes use of concepts which are, at best, poorly understood. Briefly stated, stimulus change of the sort occasioned by stimulus or task heterogeneity may reduce the monotony of the task. At a physiological level, this may be equivalent to suggesting that stimulus change increases the general activation level of the nervous system. In any event, such a motivational effect would be most apparent in the latency data, which expectation is borne out in both Exp. II and Exp. III findings. A second point in favor of the motivation hypothesis is the fact that stimulus generalization, upon which the alternative hypothesis is based, could hardly be expected to be very strong in the present situation, even under homogeneous conditions. Indeed, stimuli were chosen for ease of discrimination, and were then differentiated further by six training trials. Certainly, whatever generalization gradients remained would be extremely peaked. It would appear, therefore, that the influence of stimulus heterogeneity on performance is predominantly a motivational one.

Two additional results are of interest in that they bear upon the findings of earlier investigations. First, the visual stimulus mode was superior to the auditory mode with respect to the rate of information transmitted. This is in clear agreement with the conclusions reached by Klemmer (1956) and Mowbray (1953). Second, the manual response mode led to significantly higher transmission rates than did the verbal mode, a result in strict disagreement with the findings of several investigators (Alluisi & Martin, 1957; Alluisi & Muller, 1958). One explanation of this discrepancy has been stated earlier: the present coding scheme favored the manual mode, at least in the stimulus heterogeneity condition, through the provision of left- vs. right-hand response cues. However, superiority of the manual mode was not limited to the stimulus heterogeneity condition; it occurred under the homogeneous control tasks of Exp. I as well. Any cuing available under the latter conditions would also characterize the Alluisi research, since both were similar in the assignment of stimuli to the two hands. The discrepancy in results must therefore be attributed to those features upon which the studies differed, namely, amount of stimulus and response information and S-R compatibility. In regard to the latter variable, it is unfortunate that differences were not obtained between the various subtasks with respect to S-R compatibility. Had such differences appeared, it would have been meaningful to evaluate the influence of S-R compatibility upon the obtained heterogeneity effects.

In conclusion, then, it appears that task performance may be enhanced by the presence, in certain amounts, of stimulus or task heterogeneity. It is to be expected that the heterogeneity effect is not unlimited: the most effective level probably depends upon other task variables such as pacing mode, speed stress, stimulus entropy, S-R compatibility, level of practice, and the like. However, the present series of experiments serve to emphasize that in selecting task conditions the major concern

should be with finding *optimum* rather than *minimum* heterogeneity levels.

Summary

Three experiments were carried out to determine the influence of stimulus (S), response (R), and S-R or task heterogeneity on simple information-processing performance. Heterogeneity was provided by two dissimilar S classes (auditory vs. visual) and two dissimilar R classes (manual vs. verbal). The conditions were the types of heterogeneity noted above plus a set of homogeneous control conditions. Under all conditions there were 6 alternative Ss and 6 associated Rs; 300 serial S-R occurrences comprised each task. The major performance index was average rate of information transmitted (H_t/sec). Eighteen *Ss* served under each of the three heterogeneity conditions and 9 served under each of the four homogeneous control conditions.

The major results were: (*a*) For the particular stimulus and response classes used there appeared no evidence for S-R compatibility differences, and compatibility generally was low. (*b*) In support of earlier findings, the visual code was superior to the auditory code; contrary to earlier findings, the manual response mode was superior to the verbal response mode. (*c*) Task heterogeneity led to significantly higher rates of information transmission than did task homogeneity. (*d*) The task heterogeneity effect may be attributed partially to the S component, since S heterogeneity also enhanced performance significantly over the homogeneous controls, whereas R heterogeneity did not. Neither S nor R heterogeneity scores differed significantly from task heterogeneity scores.

From these findings it is concluded that simple information-processing performance may be enhanced by the presence, in certain amounts, of stimulus or task heterogeneity. The most effective heterogeneity level for any particular task, however, is probably determined by a number of other task variables.

References

Alluisi, E. A. Computational formulae for a distribution-free test of analysis of variance hypotheses. *USAF WADC tech. Rep.*, 1956, No. 56-339.

Alluisi, E. A., & Martin, H. B. Comparative information-handling performance with symbolic and conventional arabic numerals: Verbal and motor responses. *USAF WADC tech. Rep.*, 1957, No. 57-196.

Alluisi, E. A., & Muller, P. F., Jr. Verbal and motor responses to seven visual codes: A study in S-R compatibility. *J. exp. Psychol.*, 1958, **55**, 247-254.

Bahrick, H. P., Noble, M., & Fitts, P. M. Extra-task performance as a measure of learning a primary task. *J. exp. Psychol.*, 1954, **48**, 298-302.

Baker, K. W., Wylie, R. C., & Gagné, R. M. The effects of an interfering task on the learning of a complex motor skill. *J. exp. Psychol.*, 1951, **41**, 1-9.

Briggs, G. E., & Howell, W. C. On the relative importance of time sharing at central and peripheral levels. *USN Train. Dev. Cent. tech. Rep.*, 1959, No. 508-2.

Briggs, G. E., & Waters, L. K. Training and transfer as a function of component interaction. *J. exp. Psychol.*, 1958, **53**, 379-387.

Broadbent, D. E. *Perception and communication*. New York: Pergamon, 1958.

Ghiselli, E. E., & Brown, C. W. *Personnel and industrial psychology*. New York: McGraw-Hill, 1948.

Halsey, R. M., & Chapanis, A. On the number of absolutely identifiable spectral hues. *J. Opt. Soc. Amer.*, 1951, **41**, 1051-1058.

Klemmer, E. T. Time sharing between auditory and visual channels. In G. Finch and F. Cameron (Eds.), *Air force human engineering, personnel, and training research*. (NAS-NRC Publ. No. 455). Washington: National Academy of Sciences-National Research Council, 1956. Pp. 199-203.

Klemmer, E. T. Time sharing between frequency-coded auditory and visual channels. *J. exp. Psychol.*, 1958, **55**, 229-235.

Mowbray, G. H. Simultaneous vision and audition: The comprehension of prose passages with varying levels of difficulty. *J. exp. Psychol.*, 1953, **46**, 365-372.

Woodworth, R. S., & Schlosberg, H. Experimental psychology. (Rev. ed.) New York: Holt, 1955.

(Received November 16, 1960)

Journal of Experimental Psychology
1962, Vol. 63, No. 3, 244–247

SPACE PERCEPTION AMONG UNILATERALLY PARALYZED CHILDREN AND ADOLESCENTS [1]

HOWARD T. BLANE

Massachusetts General Hospital, Boston

The sensory-tonic field theory of perception (Werner & Wapner, 1949, 1952) assumes that the degree of muscular involvement is an important variable in perceptual experience. Utilizing this assumption, previously reported studies (Carini, 1955; Goldman, 1953; Wapner, Werner, & Chandler, 1951; Wapner, Werner, & Morant, 1951; Werner, Wapner, & Chandler, 1951) demonstrate that experimentally induced changes in the muscular state of *S* affect the visual perception of verticality and the autokinetic phenomenon. Further, in a wedding of sensory-tonic with Werner's developmental theory (Werner, 1948), Wapner and Werner (1957) have shown a developmental ordering of such effects in the perception of verticality from childhood through adulthood. In keeping with the assumption of the importance of muscular involvement as well as with previously reported findings, two specific hypotheses regarding the visual organization of space may be formulated: (*a*) Enduring asymmetric muscular imbalance in the organism will affect perceptual spatial organization, and (*b*) such spatial organization will vary with developmental level.

PROCEDURE

Four groups of *S*s were selected to participate in two experiments chosen to test the hypotheses advanced above.

Subjects.—Forty male and female postpoliomyelitic adolescents and children were divided into four equal groups: (*a*) children, 9–12 yr. of age (mean = 10.3 yr.), with residual paralysis of the right leg; (*b*) children, 9–12 yr. of age (mean = 11.1 yr.), with residual paralysis of the left leg; (*c*) adolescents, 15–19 yr. of age (mean = 16.5 yr.), with residual paralysis of the right leg; and (*d*) adolescents, 15–19 yr. of age (mean = 16.5 yr.), with residual paralysis of the left leg. The *S*s were selected from a group of 400. postpoliomyelitic children and adolescents being evaluated for skeletal growth at Children's Hospital, Boston, and care was taken to include no case with known psychiatric disturbance, visual defect, or social or mental retardation. Each patient was given a muscle test rating by a physiotherapist, on a scale ranging from 0 (complete paralysis of the affected leg) to 120 (no paralysis).[2] Since scores of 100 or more reflect only barely discernible paralysis, no patient with a score over 100 was included in the final sample. Analysis of variance revealed no significant differences among the four groups with regard to degree of paralysis.

The age at which *S* contracted poliomyelitis and the period of time elapsed between onset and participation in this investigation may be assumed to be variables of consequence, especially from a developmental viewpoint, in *S*'s perceptual functioning. However, to control these variables and simultaneously control the extent of residual paralysis, on the assumption that the latter is a more critical variable, was not possible. Consequently, child and adolescent *S*s differ significantly with respect to age at onset ($F = 10.11$, $df = 1$, $P < .01$), the children being younger at onset (mean = 3.2 yr.) than the adolescents (mean = 5.9 yr.). Similarly, elapsed time since onset is significant ($F = 15.35$, $df = 1$, $P < .01$), the children having had paralysis for an average

[1] This investigation was supported by USPHS Grant M-348 from the National Institute of Mental Health, United States Public Health Service, and was carried out when the author was a graduate student at Clark University. Actual testing of *S*s was done at Children's Hospital, Boston.

[2] I am indebted to Margaret Anderson and her staff at the Growth Study for their assistance in the selection and ratings of *S*s.

of 7.5 yr., the adolescents 10.6 yr. The groups show no significant difference, however, with respect to number of corrective orthopedic operations or corrective orthopedic appliances (braces, crutches, canes).

Experiment I.—Visual perception of verticality under conditions of body tilt and starting position was tested first. The apparatus used, an adjustable luminescent rod and a tilting chair, as well as the physical conditions of testing, have been described by Werner, Wapner, and Chandler (1951). Two variables were systematically studied: body tilt (B)—30° left (L), erect (E), 30° right (R); and starting position of the rod (S)—30° L, 10° L, 10° R, 30° R. All combinations of these variables were employed, making a total of 12 conditions: (*a*) B, 30° R-S, 30° L; (*b*) B, 30° R-S, 10° L; (*c*) B, 30° R-S, 10° R; (*d*) B, 30° R-S, 30° R; (*e*) B, E-S, 30° L; (*f*) B, E-S, 10° L; (*g*) B, E-S, 10° R; (*h*) B, E-S, 30° R; (*i*) B, 30° L-S, 30° L; (*j*) B, 30° L-S, 10° L; (*k*) B, 30° L-S, 10° R; (*l*) B, 30° L-S, 30° R. Each *S* in the four groups was tested under each of the 12 conditions, which were presented in random order. Under all conditions of testing, *S* instructed *E* as to how to move the rod until it appeared vertical to *S*.

The measure employed was the deviation in half degree of tilt from objective, or plumb line vertical. An angular position to the right of objective vertical (as viewed by *S*) was arbitrarily designated as +, and to the left of objective vertical, as −. The data were treated by analysis of variance designed to evaluate independent groups with repeated measurements.

Experiment II.—The autokinetic phenomenon was the second means chosen to test the above hypotheses. The stimulus object was a pinpoint source of light placed 10 ft. directly in front of the seated *S*. Movement of *S* was minimized by seating him in a chair with side supports and an adjustable head rest. The task for *S* was to report when the light began to move, the direction of movement, and any change of direction of movement. The *S* reported for 35 sec. after he first perceived movement. Three electric timers, activated by microswitches operated by *E*, recorded *S*'s reports of movement (*a*) to the left, (*b*) to the right, or (*c*) up or down. Combined movement reports, such as "right-up," were recorded by activation of two timers. A fourth timer automatically stopped the recording timers at the end of the 35-sec. interval. The measure used for statistical analysis was the percentage of movement reported to the right or to the left out of total movement reported, expressed as follows:

$$X = \frac{t_1 - t_2}{t_1 + t_2 + t_3} \times 100$$

where t_1 = time, in seconds, that right movement was reported; t_2 = time that left movement was reported; and t_3 = time that up-down movement was reported. Thus, a positive percentage indicates predominant movement reported to the right, a negative percentage, movement to the left. The data were treated by a double classification analysis of variance.

Results

Experiment I.—The analysis of variance is summarized in Table 1. The *F* test for the interaction between age and paralysis independent of body tilt and starting position is significant. The mean positions of

TABLE 1

Analysis of Variance of Effect of Body Tilt and Starting Position on the Position of the Apparent Vertical

Source of Variance	*df*		*MS*	*F*
Between *Ss* (*S*)	39			
Between ages (A)		1	35.00[a]	1.62
Between paralyses (P)		1	4.00[a]	0.19
A × P		1	109.00[a]	5.05*
Pooled *Ss* × (A × P)		36	21.58	
Within *Ss*	440			
Between body tilt & starting position	11			
Between body tilt (B)		2	565.00[b]	26.43**
Between starting position (S)		3	1307.00[c]	88.73**
B × S		6	80.50[d]	10.45**
S × B × S	429			
B × A		2	4.00[b]	0.19
B × P		2	54.50[b]	2.55
S × A		3	97.33[c]	6.61**
S × P		3	31.67[c]	2.15
B × S × A		6	2.17[d]	0.28
B × S × P		6	13.33[d]	1.73
B × A × P		2	57.50[b]	2.69
S × A × P		3	4.00[c]	0.27
B × S × A × P		6	12.33[d]	1.60
Pooled *Ss* × (B × S)	396			
Combinations				
Pooled *Ss* × B		72	21.38	
Pooled *Ss* × S		108	14.73	
Pooled *Ss* × (B × S)		216	7.70	
Total	479			

Note.—Degrees corrected for sign.
[a] Tested against Pooled *Ss* × (A × P).
[b] Tested against Pooled *Ss* × B.
[c] Tested against Pooled *Ss* × S.
[d] Tested against Pooled *Ss* × (B × S).
* $P < .05$.
** $P < .01$.

TABLE 2

MEAN POSITIONS OF APPARENT VERTICAL (IN DEGREES) AS A FUNCTION OF BODY TILT, STARTING POSITIONS, AND *Ss*

Degrees	Body Tilt	Starting Position		
		Children	Adolescents	All *Ss*
30° L	+2.0	−4.1	−2.8	−3.4
10° L		−2.4	−1.8	−2.1
0°	−0.4			
10° R		+2.3	+1.0	+1.8
30° R	−1.7	+4.9	+2.4	+3.6

Note.—A plus indicates a position to *Ss'* right, a minus sign to the left of the objective vertical.

apparent vertical for children with right and left paralysis were +0.6 and −0.1, respectively. The comparable means for right and left paralyzed adolescents were −0.9 and +0.3. Between the children, the systematic deviation of apparent vertical is to the right in right paralytics, and to the left in left paralytics; between the adolescents, on the other hand, apparent vertical is to the left in right paralytics and to the right in left paralytics.

The *F* tests for body tilt, starting position, and starting position by age are also significant. The mean positions for each of these conditions are presented in Table 2. With respect to body tilt, apparent vertical is located relatively opposite to the side of tilt; that is, with body tilt right the apparent vertical is located relatively left, and vice versa. The means for starting position show that apparent vertical is located relatively close to the position at which the rod was started. That this effect is greater for children than adolescents is evident in the means for starting position by age.

Experiment II.—Table 3 presents a summary of the analysis of variance of the percentages after correction for sign. The *F* test for paralysis is significant, such that right paralytics, regardless of age, report predominant movement to the left (mean = −29.0%), whereas left paralytics report predominant movement relatively to the right (mean = +2.6).

DISCUSSION

This study demonstrates that visual perception of space as reflected in the experience of verticality and apparent movement varies with enduring asymmetric muscular imbalance, such that right paralyzed children and adolescents locate apparent vertical and perceive apparent motion differently from left paralyzed children and adolescents. Furthermore, with respect to verticality, a developmental shift between children and adolescents is evident. These results further confirm the assumption of sensory-tonic field theory with regard to the importance of muscular involvement in perceptual experience. The present findings are also consistent with those of previous experiments which show that extraneous stimulation influences perception (Carini, 1955; Goldman, 1953; Wapner, Werner, & Chandler, 1951; Wapner, Werner, & Morant, 1951; Werner et al., 1951), and indicate that unilateral paralysis may be thought of as one form of extraneous stimulation. The developmental shift between children and adolescents in the experiment on apparent verticality is congruent with developmental studies of perception in

TABLE 3

ANALYSIS OF VARIANCE OF DIRECTION OF APPARENT MOTION IN PARALYZED CHILDREN AND ADOLESCENTS

Source of Variance	*df*		*F*
Between ages (A)		1	1.00
Between paralyses (P)		1	5.36*
A × P		1	0.14
Within *Ss MS*		36	(1864.13)
Total	39		

* $P < .05$.

normal *Ss* reported by Wapner and Werner (1957). That no developmental shift occurred in the experiment on apparent movement may relate to the fact that the autokinetic situation, being minimally dimensional, does not permit the subject-object differentiation assumed to be necessary for such a shift to occur.

Summary

Two experiments were carried out to investigate the influence of asymmetric muscular imbalance in children and adolescents on the location of apparent vertical and on the laterality of apparent movement. Four groups of 10 *Ss* each were tested. Two groups consisted of children, one with right leg paralysis, the other with left leg paralysis. The remaining two groups were composed of adolescents, one with right leg paralysis, the other with left leg paralysis.

Experiment I showed that adolescent paralytics locate the apparent vertical opposite to the side of paralysis, whereas paralytic children locate the apparent vertical to the side of paralysis. Experiment II showed that paralyzed children and adolescents report apparent movement predominantly away from the side of paralysis. These results, viewed within the framework of sensory-tonic field theory, support the assumption that muscular involvement affects perceptual experience.

References

Carini, L. P. An experimental investigation of perceptual behavior in schizophrenia. Unpublished doctoral dissertation, Clark University, 1955.

Goldman, A. E. Studies in vicariousness: Degree of motor activity and the autokinetic phenomenon. *Amer. J. Psychol.*, 1953, **66**, 613–617.

Wapner, S., & Werner, H. *Perceptual development.* Worcester: Clark Univer. Press, 1957.

Wapner, S., Werner, H., & Chandler, K. A. Experiments on sensory-tonic field theory of perception: I. Effect of extraneous stimulation on the visual perception of verticality. *J. exp. Psychol.*, 1951, **42**, 341–345.

Wapner, S., Werner, H., & Morant, R. B. Experiments on sensory-tonic field theory of perception: III. Effect of body rotation on the visual perception of verticality. *J. exp. Psychol.*, 1951, **42**, 351–357.

Werner, H. *Comparative psychology of mental development.* Chicago: Follett, 1948.

Werner, H., & Wapner, S. Sensory-tonic field theory of perception. *J. Pers.*, 1949, **18**, 88–107.

Werner, H., & Wapner, S. Toward a general theory of perception. *Psychol. Rev.*, 1952, **59**, 324–338.

Werner, H., Wapner, S., & Chandler, K. A. Experiments on sensory-tonic field theory of perception: II. Effect of supported and unsupported tilt of the body on the visual perception of verticality. *J. exp. Psychol.*, 1951, **42**, 346–350.

(Received December 6, 1960)

Journal of Experimental Psychology
1962, Vol. 63, No. 3, 248–253

PERFORMANCE IN INSTRUMENTAL CONDITIONING AS A JOINT FUNCTION OF TIME OF DEPRIVATION AND SUCROSE CONCENTRATION[1]

JOHN R. STABLER[2]

University of Texas

As Pubols (1960) has pointed out in a recent review of the literature, there have been few experiments reported dealing with the interaction of incentive-magnitude with other variables. In particular, more research is needed in which both drive (D) and reward (K) are varied at the same time before a reasonably complete specification of the laws relating D and K can be formulated.

Ramond (1954) compared the performance of white rats in a simple runway as a function of variation in time of deprivation (T_d) and delay of reinforcement (T_g). Experiments by Fredenburg (1956) and Besch and Reynolds (1958) involved the simultaneous manipulation of T_d and length of the response chain (L_g), i.e., distance that S had to run to reach the goal. The data from these experiments indicated no interaction between D and K; however manipulating either T_g or L_g introduces variable opportunity for competing responses to occur. A fatigue variable may also be differentially present when L_g is varied. The results from studies which have manipulated the amount of the incentive (W_G) are contradictory (Pubols, 1960). One reasonable explanation for the inconclusiveness of the results from such studies is suggested by Guttman's (1953) analysis of this particular method of varying reward magnitude. He points out that a variation in the size or number of the incentive object introduces several simultaneous sources of variation. A large incentive object has different visual, olfactory, and tactile stimulus features and requires more chewing and swallowing than a small incentive object. The purpose of the present study was to determine the nature of the relationship between D and K when different sucrose concentrations are used to vary K, thereby controlling for the possibility of differences in preconsummatory stimulation, ingestive activity, and presence of competing responses.

A second purpose of the present experiment was to provide additional information on the dual reinforcing effect of sucrose solutions. Previous experimentation (Wetzel, 1959) has indicated that rats inexperienced with sucrose respond to it on the basis of its palatibility. On the other hand, if hungry Ss are fed adequate quantities of sucrose solutions over a period of days, they will subsequently respond to the solution not only for the sensory stimulation that it provides but also for its hunger-reducing potential. By employing a within-groups instead of a between-groups design, the present experiment provides evidence as to whether or not the second reinforcing property of sucrose becomes manifest over a prolonged period of small-quantity feedings.

[1] Part of a dissertation submitted to the Graduate School of the University of Texas in partial fulfillment of the requirements for the PhD degree. The author is indebted to Alfred Castaneda for advise and assistance throughout the course of the investigation.

[2] Now at Louisiana State University.

METHOD

Subjects and apparatus.—The *S*s were 90 experimentally naive male rats of the Wistar strain purchased from Harlan Industries, Cumberland, Indiana. All *S*s were 85–90 days old when purchased and 106-111 days old at the beginning of pretraining.

The apparatus was a straight runway consisting of a 12-in. start box, a 24-in. alley, and a 12-in. goal box. The runway was 6 in. high and 3 in. wide. It was covered by wire mesh. All inside spaces were painted flat black. A 100-w. bulb was located 55 in. above each end of the alley. The light from these bulbs was shaded by a 60 × 12 in. sheet of fiberboard. The start box and alley were separated by two guillotine doors, one of clear plastic and the other of opaque fiberboard.

Running-time measures were provided by interruption of monochromatic red beams to photocells which controlled relays and standard timers. The raising of the plastic door started the first clock. Interruption of the first beam, 6 in. into the alley, stopped the first clock and started a second. Interruption of a second beam, located 30 in. into the alley, stopped the second clock and started a third, which stopped when a beam located 11 in. into the goal box was interrupted. The final beam was located directly over the food cup.

Experimental design.—In Stage 1 of training, two levels of hunger (4 hr. and 20 hr.) and three levels of sucrose concentration (8%, 16%, and 32%) were employed in a 2 × 3 factorial design. The *S*s were randomly assigned to the six experimental conditions designated as: D_4K_8, D_4K_{16}, D_4K_{32}, $D_{20}K_8$, $D_{20}K_{16}$, and $D_{20}K_{32}$. In each of two groups 1 *S* died, and in a third group, 2 *S*s died. To equate the number of *S*s within groups, 1 randomly selected *S* was discarded from each of the three remaining groups, and the missing data for 1 *S*, which died after the completion of Stage 1, were prorated from the scores of *S*s remaining within its group. This left 14 *S*s in each group, 84 *S*s in all. The three sucrose concentrations were designated according to percentage of concentration by weight. Thus, the 8% solution was prepared by combining 8 gm. of sucrose (commercial table sugar) and 92 cc of distilled water. The solutions were kept in closed containers and were prepared fresh before each day of testing.

In Stage 2, each group was divided into matched experimental and control groups of 7 *S*s each. This division resulted in a 2 × 2 × 3 factorial design. Fifteen minutes after each testing session the experimental *S*s were allowed 15 min. access to sugar solution of the same percentage as that concentration whicht hey had consumed in the goal box. The control *S*s continued under the same conditions as in the first stage. All *S*s were fed 2 hr. after the time of measurement.

Procedure.—Two weeks prior to the first training day, *S*s were placed on their respective food deprivation schedules. Water was continuously available except during the experimental sessions. During the 10 days prior to the first training day, *S*s were adapted to handling and to the apparatus. The adaptation procedure closely approximated the procedure to be followed during the experiment.

Each *S* was given three trials a day on each of the 20 days of Stage 1 and each of the 20 days of Stage 2. The following procedures were employed on every trial for each *S*: *S* was placed in the start box, and the opaque door was opened as soon as *S* oriented to it. The plastic door was removed ½ sec. later. As soon as the two drops of incentive substance had been consumed, *S* was removed and placed in a waiting cage until the next *S* had run a trial, when it was returned to its home cage for the remainder of the intertrial interval. Intertrial intervals were approximately 20 min. long. The number of times that *S* stopped its forward movement in each alley segment and the time taken to traverse each segment were recorded after every trial.

RESULTS

Running speed.—The final performance levels attained under the six experimental conditions were of primary interest; hence the mean running speeds (reciprocal time in ft/sec) over Trials 31–60 were computed for each *S*. The group mean running speeds in the start box and in the runway for Trials 31–60 are plotted in Fig. 1 and 2. Because the results of the analyses of variance of the mean goal box speeds were comparable to those of the start box and runway, i.e., all showed a significant interaction, the data based on that particular segment were not further analyzed. Since previous research (Besch & Reynolds, 1958) has suggested that

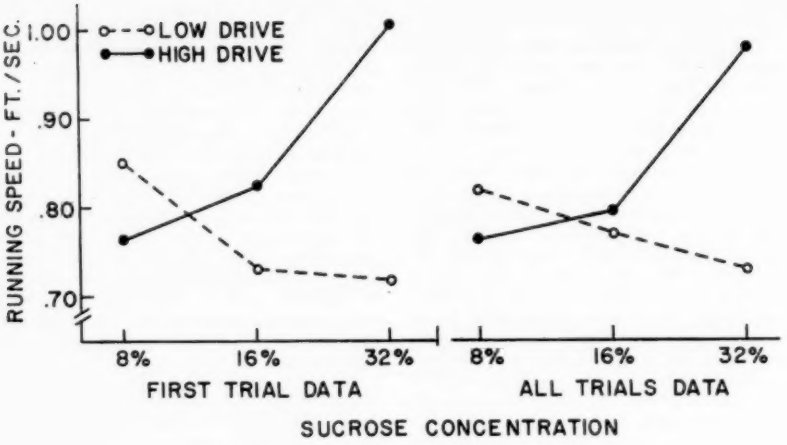

FIG 1. Mean start box speeds for Trials 31–60.

data based on the first trial of each day, rather than on all trials of each day, may differ with respect to amount of *D* and *K* interaction, both types of data were analyzed in the present study. Start box and runway measures and measures based upon either the first trial of each day or upon all trials of each day were generally consistent. There was, however, some evidence of a "warm up" effect from the first trial of each day. Runway speed measures based on all trials of each day were higher than those based on only the first trial of each day.

A significant interaction was found between the effects of drive and

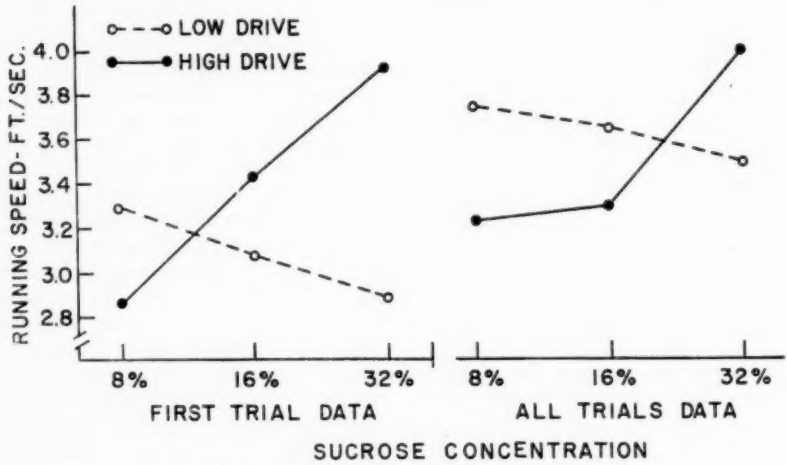

FIG. 2. Mean runway speeds for Trials 31–60.

incentive with each of the four sets of data ($F = 4.58$, $P < .05$; $F = 4.94$, $P < .01$; $F = 4.19$, $P < .05$; $F = 3.15$, $P < .05$; respectively, with $df = 2/78$ in each case). Trend analyses (Grant, 1956) indicated that the variance associated with interaction had a highly significant linear component (Fs = 9.16, 9.21, 8.38, and 6.11, respectively, with $P < .01$, $df = 1/78$ in all cases). Duncan multiple range tests (MRT) were performed in order to make orthogonal comparisons between the six experimental groups. The four tests consistently showed that Group $D_{20}K_{32}$ was significantly faster than Group $D_{20}K_8$. This difference was the only significant one obtained from the mean runway speeds based on all trials of each day. According to tests of start box and runway speeds based on the first trial of each day, Group $D_{20}K_{32}$ was also significantly faster than either Group D_4K_{16} or Group D_4K_{32}. The MRT performed on the start box speeds based on all trials of each day showed a significant difference between Group $D_{20}K_{32}$ and each of the others. Although the graphs indicate an inversion of drive groups at K_8, none of the MRTs showed a significant difference between Groups D_4K_8 and $D_{20}K_8$. It should also be noted that in no case were the means of Groups D_4K_8 and D_4K_{32} significantly different from each other.

TABLE 1

NUMBER OF STOPS FOR TRIALS 31–60, BASED ON ALL TRIALS FOR EACH DAY

Group	Alley Segment					
	Start Box		Runway		Goal Box	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
D_4K_8	3.1	3.13	2.8	2.88	3.1	3.84
D_4K_{16}	4.9	4.02	4.0	2.95	3.1	3.55
D_4K_{32}	5.9	3.50	2.1	1.35	1.4	2.09
$D_{20}K_8$	5.3	4.78	3.7	2.94	3.1	3.57
$D_{20}K_{16}$	4.3	4.48	2.1	2.22	1.5	1.63
$D_{20}K_{32}$	1.4	1.84	.7	1.22	.5	1.05

TABLE 2

MEAN RUNNING SPEEDS FOR TRIALS 91–120, BASED ON ALL TRIALS FOR EACH DAY

Group	Alley Segment			
	Start Box		Runway	
	Exp.	Control	Exp.	Control
D_4K_8	.94	1.00	4.03	3.76
D_4K_{16}	.89	.86	3.78	3.32
D_4K_{32}	.88	.95	3.26	3.37
$D_{20}K_8$	.96	1.04	3.72	3.67
$D_{20}K_{16}$	1.08	.92	3.84	3.74
$D_{20}K_{32}$	1.11	1.18	4.19	4.46

Stop behavior.—The mean numbers of stops for each group in the start box, the runway, and the goal box for Trials 31–60, based on all trials of each day, are presented in Table 1. A significant interaction was obtained from the analysis of variance of the number of stops in the start box ($F = 5.03$; $P < .01$, $df = 2/78$); however the runway and the goal box analyses showed only a significant incentive effect (Fs = 4.06 and 3.79, respectively, with $P < .05$, $df = 2/78$ in each case). The results of the MRTs suggested that the loss of the interaction and appearance of an incentive effect in the runway and goal box was a function of the change in relative position of Group D_4K_{32}. The significant difference which was found in running speeds between Groups $D_{20}K_{32}$ and $D_{20}K_8$ occurred also for stop behavior. Also, as with the running speed data, no significant differences were obtained between Groups D_4K_8 and $D_{20}K_8$, or between Groups D_4K_8 and D_4K_{32}. The correlations between the two dependent variables for start box, runway, and

goal box, based on all trials of each day, were −.72, −.63, and −.42, respectively.

Results of Stage 2.—The mean running speeds for each group in the start box and the runway for Trials 91–120, based on all trials of each day, are presented in Table 2. Analyses of the data showed a reliable D and K interaction effect for runway speeds ($F = 5.66$, $P < .01$, $df = 2/72$) and a reliable D effect for both start box and runway speeds ($F = 13.09$, $P < .01$; $F = 5.38$, $P < .05$; respectively, with $df = 1/72$ in each case). It may be recalled that no significant D effect was obtained from the analyses of Trials 31–60. Multiple range tests performed on start box and runway data showed no significant differences between experimental and control groups at any given D and K combination of conditions. For example, although the experimental and control groups of combination $D_{20}K_{32}$ ran significantly faster than several other groups, they did not differ significantly from each other. The groups representing the various D and K combinations held the same relative positions at the end of Stage 2 as they had held at the end of Stage 1. The start box and runway data for Trials 91–120, based on all trials of each day, were also covaried upon the individual means of Trials 31–60 in order to control statistically for the level of performance that had been attained before the beginning of Stage 2. The D effect was found to be highly reliable ($F = 10.13$, $F = 13.10$; $P < .01$, $df = 1/71$ in each case).

Discussion

Relation of D *and* K *variables.*—The evidence from this experiment is clearly more compatible with Hull's position that drives and incentives are multiplicatively related than with Spence's proposal that drives and incentives join in an additive fashion. One implication of the results is that the method used to vary the amount of reinforcement, i.e., L_g, T_g, W_G, etc., is of critical importance to theoretical predictions regarding incentive effects.

With the exception of Group D_4K_{32}, the stop data appeared to be inversely related to running speed measures. This is consistent with the observations of Cotton (1953) and others that running speed measures are partly dependent upon the number of times S stops in the alley or performs some other nonadjustive response.

Hypotheses about incentive action.—Spence (1956) has proposed that one determinant of the strength of an incentive is the generalized habit strength (which here will be denoted $\bar{H}_K$) between alley stimuli and anticipatory consummatory behavior ($r_g - s_g$). Several recent eyelid conditioning experiments (Spence, Haggard, & Ross, 1958; Trapold & Spence, 1960) have indicated that the magnitude of H varies as a function of the intensity of the unconditioned stimulus (UCS). In such respects as the abruptness of the change in sensory stimulation which it produces, using a sucrose solution as the UCS appears analogous to an air-puff UCS; thus incentive magnitude in the present study may have been a function of differential strength of $\bar{H}_K$. If the equation for excitatory potential (E) were to be written $E = (H + \bar{H}_K) \times D$, it would be possible to derive that D and K combine multiplicatively as suggested by the present experiment.

$$E_{\text{High}} = (H + \bar{H}_{K_{\text{High}}}) \times D$$
$$E_{\text{Low}} = (H + \bar{H}_{K_{\text{Low}}}) \times D$$

Subtracting:

$$E_{\text{High}} - E_{\text{Low}} = D(\bar{H}_{K_{\text{High}}} - \bar{H}_{K_{\text{Low}}})$$

Dual reinforcing effect of sucrose solutions.—The finding of a significant D effect at the end of Stage 2, together with the absence of a significant D effect at the end of Stage 1, indicates that the effect of level of deprivation upon run-

ning speed is greater if *Ss* have been rewarded with small amounts of sucrose solution over an extended period of time. It suggests that *Ss* learned to respond to the incentive in terms of its need-reducing potential as well as in terms of its palatability. The absence of reliable experimental versus control main effects and interaction effects indicates that the *D* effect was determined by experience with sucrose rewards on the 60 trials of Stage 2 rather than by extra sucrose consumption in the home cages.

Summary

The experiment was primarily concerned with the joint effects of concentration of sucrose incentive and hours of food deprivation upon locomotor behavior in a straight alley. Three levels of concentration and two levels of drive were employed in a factorial design with six groups of male rats. Running speed and frequency of stops were recorded for 120 trials, i.e., 60 trials in Stage 1 and 60 trials in Stage 2.

The major finding from Stage 1 was that asymptotic running speed measures showed a significant interaction between drive and incentive variables. Frequency of stops was generally inversely related to running speed. Start box and runway measures, and measures based upon either the first trial of each day or upon all trials of each day, were generally consistent. The results were discussed in terms of Spence's (1956) incentive motivation hypotheses concerning the action of incentives in instrumental reward conditioning.

In Stage 2 half the *Ss* in each experimental group were allowed to consume large quantities of the same concentration sucrose solution as they had been receiving in the goal box. The solutions were provided after each day's testing and 90 min. before feeding. A significant *D* effect, obtained from analyses of the asymptotic running speeds at the end of the second stage, was interpreted as evidence that after a prolonged period of small-quantity feedings of sucrose solution *Ss* came to respond to the incentive in terms of its need-reducing potential as well as in terms of its palatability.

References

Besch, N. F., & Reynolds, W. F. Alley length and time of food deprivation in instrumental reward learning. *J. exp. Psychol.*, 1958, **56**, 448–452.

Cotton, J. W. Running time as a function of amount of food deprivation. *J. exp. Psychol.*, 1953, **46**, 188–198.

Fredenburg, N. C. Response strength as a function of alley length and time of deprivation. *Proc. Ia. Acad. Sci.*, 1956, **63**, 563–567.

Grant, D. A. Analysis of variance in the analysis and comparison of curves. *Psychol. Bull.*, 1956, **53**, 141–155.

Guttman, N. Operant conditioning, extinction, and periodic reinforcement in relation to concentration of sucrose used as reinforcing agent. *J. exp. Psychol.*, 1953, **46**, 213–224.

Pubols, B. H. Incentive magnitude, learning, and performance in animals. *Psychol. Bull.*, 1960, **57**, 89–115.

Ramond, C. K. Performance in instrumental learning as a joint function of delay of reinforcement and time of deprivation. *J. exp. Psychol.*, 1954, **47**, 248–251.

Spence, K. W. *Behavior theory and conditioning.* New Haven: Yale Univer. Press, 1956.

Spence, K. W., Haggard, D. F., & Ross, L. E. UCS intensity and associative (habit) strength of the eyelid CR. *J. exp. Psychol.*, 1958, **55**, 404–411.

Trapold, M. A., & Spence, K. W. Performance changes in eyelid conditioning as related to the motivational and reinforcing properties of UCS. *J. exp. Psychol.*, 1960, **59**, 209–214.

Wetzel, R. J. The effect of experience with a taste reward. *J. comp. physiol. Psychol.*, 1959, **52**, 267–272.

(Received December 9, 1960)

Journal of Experimental Psychology
1962, Vol. 63, No. 3, 254–257

CONCEPT IDENTIFICATION OF AUDITORY STIMULI AS A FUNCTION OF AMOUNT OF RELEVANT AND IRRELEVANT INFORMATION [1]

ROSARIA G. BULGARELLA [2] AND E. JAMES ARCHER

University of Wisconsin

A number of experiments (Archer, Bourne, & Brown, 1955; Bourne, 1957; Brown & Archer, 1956) have studied the effect of variations in the amount of nonredundant irrelevant information on concept identification. In general, a linear degradation of performance is produced by an increase in irrelevant information.

Two additional studies (Walker, 1957; Wargo, 1960) have investigated the concurrent effect of variation of nonredundant relevant information. In both of these studies, task difficulty increased as a function of an increase in relevant and irrelevant information. However, Walker found the increase to be nonlinear and also observed a significant interaction between the two variables. The failure of Wargo to find either of these effects was assumed to be due to certain metholological differences which will be considered later.

Except for a study by Lordahl (1959), all of the previous concept identification studies have used visual stimuli only. In the exception, Lordahl presented problems which required *S* to use 1 bit of relevant visual information and 1 bit of relevant auditory information while *E* simultaneously varied the amount of irrelevant information to both sense modalities. Whereas the effect of such variation was as expected when the irrelevant information was visual, the auditory irrelevant information did not have a significant adverse effect on performance.

The purpose of the present study was to determine whether, in general, auditory information was comparable to visual information in concept identification, and in particular to determine the effect of simultaneous variation of both relevant and irrelevant *auditory* information.

PROCEDURE

Subjects.—Ninety women students from introductory and intermediate courses in psychology at the University of Wisconsin served individually for one session each. All were volunteers who had not previously served in a concept identification study.

Design.—The major independent variables were the amounts of relevant and irrelevant auditory information. A 3×3 orthogonal design with 1, 2, or 3 bits of relevant information and 0, 1, or 2 bits of irrelevant information was used.

Five different specific combinations of relevant information defined the problems at each level of relevant information, giving 15 problems in all. Since the same problems were presented at each of the three levels of irrelevant information, there were 45 different combinations of problems and irrelevant information. Two *S*s were randomly assigned to each of these 45 combinations.

Apparatus.—The auditory stimuli were presented through a pair of earphones worn by *S* who sat before a control console. The console had two lever-action switches and two lamps. Above one switch a black plus sign was painted. This switch was to be pressed for all positive instances of the concept. The other switch was unmarked, and

[1] This research was supported by a grant from the Research Committe from funds provided by the Wisconsin Alumni Research Foundation. The paper was prepared while the second-named author was a Senior Postdoctoral Fellow (NSF) at the Applied Psychology Research Unit, Cambridge, England.

[2] Now at Michigan State University.

it was to be pressed for all negative instances of the concept. When *S* made a correct response, a white lamp came *on*; when *S* made an incorrect response, a red lamp came *on*. In both cases the feedback signal continued until the next auditory pattern was presented 4 sec. later.

The apparatus which generated the auditory stimuli has been described in a previous study by Lordahl (1959). The *S* saw only the control console; the remainder of the apparatus was concealed by a partition. After giving the instructions, *E* sat behind the partition and was not visible to *S*. An Esterline-Angus 20-channel operations recorder was used to obtain a permanent record of *S*'s responses. A counter recorded *S*'s errors, and a stopwatch was used to measure the time to criterion.

Instructions and task.—The instructions were tape recorded and played to *S* at the beginning of the session. The *S* was informed that she would hear various sounds which were to be placed in the categories represented by the switches before her. She was told that those sounds which were to be placed in the category represented by the switch with a plus sign had something in common with each other, and all other sounds belonged in the unmarked category. The significance of the feedback lamps was explained and *S* was instructed to press only one switch for each tone pattern. Accuracy was stressed over speed.

The task was self-paced (*S* could take as long as she liked to respond to a tone pattern), but the next stimulus was automatically presented 4 sec. after the response was made. The *S* served until she had made either 32 consecutively correct responses or had heard a total of 192 stimuli, whichever came first. If *S* did not reach the performance criterion in 96 patterns, she was given a 1-min. rest before the session was continued.

Stimuli and problems.—The auditory stimuli could be described in terms of five bileveled dimensions. The *frequency* of the tone was either 250 cps or 1800 cps. Its *intensity* was either 43 or 64 phons for the 250-cps tone, or it was 54 or 64 phons for the 1800-cps tone. The intermittent addition of *white noise* (at 55 phons) or its absence was another dimension. The sound also had a quality of *steadiness* or regular interruption (on for 1 sec. and off for .5 sec.). The fifth dimension was the *earphone*, either left or right, through which the sound was presented.

When variation of a dimension was not needed, as in some problems where the dimension was neither relevant nor irrelevant, one level was consistently presented. The constant levels were 1800 cps for frequency, 43 or 54 phons (depending on which of the frequencies was used) for the intensity, the absence of white noise, steadiness of the tone, and the left earphone. The specific levels of the dimensions were determined by a pilot study which required comparable *S*s to show 100% discriminability of the levels.[3]

The 15 problems were of three types. Five problems had only 1 bit of information relevant. For example, when frequency was the only relevant information *S* would hear soft, clean, steady tones in the left earphone. If the tone had a high frequency, *S* was to press the switch with the plus sign. If the tone had a low frequency, *S* was to press the switch which was unmarked.

Five other problems had 2 bits of relevant information. For example, if frequency and intensity were the relevant dimensions, *S* would hear a clean, steady tone in the left earphone. If the tone had both a high frequency and a low intensity, *S* was to press the switch with the plus sign. For all other tones *S* was to press the unmarked switch.

In both of these examples no irrelevant information was present. If, however, earphone had been an irrelevant dimension, then although the above described tones would have come into each ear equally often, *S* should ignore this dimension and respond only to frequency and intensity in the latter problem, and only frequency in the former.

The third set of five problems had 3 bits of relevant information. For example, a positive instance could be a high frequency, low intensity, and interrupted tone. All others would be negative instances.

The program, which was controlled by Western Union tape transmitters, contained a series of 96 stimuli. If the concept was not identified in the first presentation of 96 patterns, the series was repeated.

The order of presentation of the negative and positive instances of the concept was determined randomly, but no matter how much information was relevant or irrelevant, 25% of the presented patterns were positive instances.

RESULTS

Trials to criterion.—As the amount of relevant information increased, the number of trials (auditory patterns) increased. For 1, 2, and 3 bits of

[3] The writers are indebted to Willard R. Thurlow for his help in assessing the frequencies and intensities of the signals.

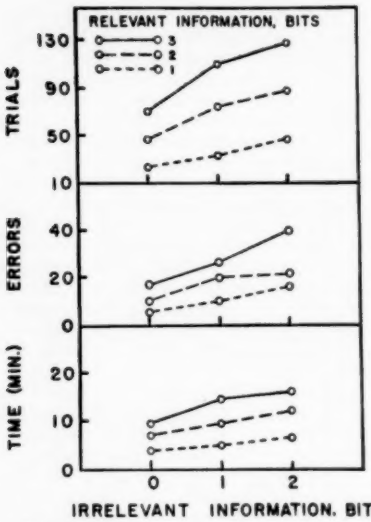

FIG. 1. Performance as measured by mean number of trials, mean number of errors, and mean time to criterion as a function of amount of nonredundant relevant and irrelevant information.

relevant information, the mean number of trials to criterion was 62.6, 99.3, and 131.4, respectively. The few *S*s who failed to achieve the performance criterion were assigned the score of 192. Failing to take all *S*s to the performance criterion introduces a slight bias in estimate for high levels of both relevant and irrelevant information. On the other hand, replacement of these less able *S*s with others who could solve the problem would have introduced a bias in the opposite direction. Figure 1 shows the mean number of trials to criterion as a function of both relevant and irrelevant information. The results of an analysis of variance indicated that the increase in trials to criterion was highly significant ($F = 23.48, df = 2/12, P < .001$) and that this function is best described by a straight line since only the linear component was significant.

For 0, 1, and 2 bits of irrelevant information, *S*s required 74.8, 103.1, and 112.0 trials, respectively, to achieve the criterion. This increase was significant ($F = 9.18$, $df = 2/45$, $P < .001$). In this case too, only the linear component was significant; the increasing performance degradation is well described by a straight line.

Of particular interest is the lack of an interaction between relevant and irrelevant information. At least over the range studied, adding irrelevant information into an already difficult auditory display did not make the task differentially more difficult.

Errors.—This measure is the number of stimuli incorrectly assigned to a category by *S*. Figure 1 shows the mean number of errors as a function of relevant and irrelevant information. For 1, 2, and 3 bits of relevant information, the mean number of errors was 9.3, 17.5, and 27.6, respectively. These differences were highly significant ($F=25.70$, $df=2/12$, $P<.001$). The function was best described by a straight line since only the linear component was significant.

A similar increase in number of errors was observed as a function of irrelevant information: the mean number of errors was 9.7, 18.8, and 25.8 for 0, 1, and 2 bits of irrelevant information, respectively. These differences were also highly significant ($F=8.24$, $df = 2/45$, $P < .001$), and again only the linear component was significant.

Time to criterion.—Figure 1 shows the time to criterion as a function of both relevant and irrelevant information. The results of an analysis of variance were similar to those for trials to criterion and errors. Both relevant and irrelevant information and their linear components were

significant and none of the other terms was significant.

Discussion

The results of the present study indicate that concept identification of auditory stimuli is quite comparable to concept identification of visual stimuli. As in previous studies (Bourne, 1957; Brown & Archer, 1956; Walker, 1957; Wargo, 1960) performance, as measured by trials, errors, and time to criterion, degraded as a linear function of amount of nonredundant irrelevant information. Only in the study by Archer, Bourne, and Brown (1955) was the performance degradation nonlinear. Apparently that finding was anomalous.

The present results are also consistent with those of Wargo (1960). Performance degraded as a linear function of nonredundant relevant information and the interaction between relevant and irrelevant information was not significant. In this regard the present study is also not consistent with Walker (1957). Since Walker required his *Ss* to use two, four, or eight response categories when 1, 2, or 3 bits of information were relevant, response complexity was confounded with concept complexity. The differential response complexity could account for the nonlinear performance degradation he observed; and this nonlinear function would, in turn, account for the significant interaction.

Of particular interest is the contrast of Lordahl's (1959) data with the present results. In his study, variation of irrelevant auditory information had little effect on performance when *S* had to use both visual and auditory relevant information simultaneously. Apparently, as Lordahl suggested, *Ss* either have a preference for responding to visual irrelevant information or they are better able to ignore auditory irrelevant information. However, when the concept to be identified involves only auditory signals, then variation in auditory irrelevant information is just as effective a way of manipulating the complexity of the task.

Summary

The effect upon concept identification of the simultaneous variation of nonredundant relevant and irrelevant auditory information was determined. The design had three levels (1, 2, or 3 bits) of relevant information orthogonal to three levels (0, 1, or 2 bits) of irrelevant information. There were five different problems within each of the three levels of relevant information. Two *Ss* were assigned to each of the 45 combinations of problems and information.

The *S*'s task was to place auditory stimuli into one of two categories, i.e., distinguish between positive and negative instances of conjunctive concepts. Ninety *Ss* served individually to a criterion of 32 consecutively correct responses or a maximum of 192 trials (whichever occurred first). The *Ss* received immediate information feedback about the correctness of their responses through two signal lights.

The major results were: (*a*) Performance, as measured by trials, errors, and time to criterion, degraded as a linear function of both nonredundant relevant and irrelevant information. (*b*) The interaction between the two variables of relevant and irrelevant information was not significant.

References

Archer, E. J., Bourne, L. E., Jr., & Brown, F. G. Concept identification as a function of irrelevant information and instructions. *J. exp. Psychol.*, 1955, **49**, 153–164.

Bourne, L. E., Jr. Effects of delay of information feedback and task complexity on the identification of concepts. *J. exp. Psychol.*, 1957, **54**, 201–207.

Brown, F. G., & Archer, E. J. Concept identification as a function of task complexity and distribution of practice. *J. exp. Psychol.*, 1956, **52**, 316–321.

Lordahl, D. S. Concept identification as a function of the amount of information in simultaneously presented visual and auditory signals. Unpublished doctoral dissertation, University of Wisconsin, 1959.

Walker, C. M. Concept identification as a function of amounts of relevant and irrelevant information. Unpublished doctoral dissertation, University of Utah, 1957.

Wargo, D. G. Concept identification as a function of relevant and irrelevant information. Unpublished master's thesis, University of Wisconsin, 1960.

(Received December 9, 1960)

Journal of Experimental Psychology
1962, Vol. 63, No. 3, 258–268

TRANSFER FROM VERBAL-DISCRIMINATION TO PAIRED-ASSOCIATE LEARNING [1]

WILLIAM F. BATTIG, JOHN M. WILLIAMS,[2] AND JOHN G. WILLIAMS [3]

University of Virginia

The need for a more analytic approach to verbal learning, whereby two or more phases or stages are distinguished and treated separately, is becoming increasingly apparent. Underwood and his associates have clearly demonstrated the importance in verbal paired-associate learning of distinguishing between a *response-learning* or *response-recall* phase and an *associative* phase (Underwood, Runquist, & Schulz, 1959), and have further suggested that quite different mechanisms may be involved in these two phases (Underwood & Schulz, 1960). The implications of a similar stage-wise analysis have also been profitably exploited in studies of serial learning (Hovland & Kurtz, 1952), as well as verbal-discrimination learning (Runquist & Freeman, 1960).

In view of the high order of complexity of verbal paired-associate learning, an extension of this kind of analytic approach to encompass more than two phases would seem to be a strong possibility, at least at a conceptual level. However, the usefulness of the Underwood et al. approach can be largely attributed to their success in separating response-learning and associative phases in terms of specific experimental operations involving some type of familiarization with the material to be learned prior to its use in a paired-associate task. Further analyses of paired-associate learning into separate stages will therefore be of value only to the extent that experimental procedures can be devised which permit the isolation and separate evaluation of these phases.

The process of discriminating between items of the list has been suggested previously as an important aspect of paired-associate learning (e.g., Gibson, 1940), and may represent an additional phase separable from either the response-learning or associative phases. Appropriate operations for the experimental isolation of a discrimination phase would seem to be provided through the use of a verbal-discrimination (VD) learning procedure prior to paired-associate (PA) learning. Accordingly, the first of the present experiments represented an attempt to isolate and determine the relative importance of a phase in PA learning involving the discrimination between items of a pair, as compared with familiarization and associative phases similar to those identified previously. This was accomplished by introducing controlled variations in the relationships existing between the material used in VD learning, and that subsequently employed in PA learning. Due to the unexpected nature of the results, which gave no evidence of any positive effect on PA performance from either prior discrimination or familiarization by means of a VD task, Exp. II was then undertaken to try to replicate and extend these findings concerning the nature of transfer between VD and PA learning.

[1] The research reported herein was performed pursuant to a contract with the United States Office of Education, Department of Health, Education, and Welfare.

[2] Now at College of William and Mary.

[3] Now at Carnegie Institute of Technology.

METHOD

Experiment I

Subjects.—The 36 *S*s were undergraduates who were either paid for their services or satisfied an introductory psychology course requirement. Additional motivation was provided by an announced bonus of $5 and $2 to the two fastest learners. Each *S* served for a single session of 1 hr. or less, during which he first learned a VD list, immediately followed by a PA list.

Materials.—The nonsense-syllable materials were all selected from the 34–37% range of association values as determined by Archer (1960). A single list of 12 pairs was constructed for PA learning, taking special precautions to minimize formal similarity between any of the items in the list and to avoid obvious associations within any pair. An additional eight syllables, chosen so as to be minimally similar to those on the PA list and to one another, were selected for use in the VD task, in conjunction with the syllables that were also used in the PA list.

VD procedure.—Each *S* was presented with a list of 12 pairs of nonsense-syllables, one of each pair having been previously designated as "right" and the other "wrong," and instructed to learn to respond by pronouncing the right syllable each time that pair was presented. Each pair was typed on a 3 × 5 in. card and presented by *E* through a manually operated card exposure device under self-paced conditions, until *S* had pronounced one of the two syllables on the card. At this point, *E* said either "right" or "wrong," after which the card containing the next pair was presented immediately. The use of self-paced conditions was necessary to insure that a response was made to each pair, although *S*s typically responded within 2 or 3 sec., and few deviated appreciably from a regular rate of responding. A prearranged order of presentation provided that (*a*) all 12 pairs were presented once on each trial; (*b*) no pair appeared in the same position within the list on more than one trial (unless *S* required more than 12 trials, in which case the orders used on the first 12 trials were repeated); (*c*) each syllable appeared in the left and right positions on the card equally often; (*d*) an approximately equal number of right and wrong syllables appeared in each position on the card within any trial. Learning was carried out to a criterion of two successive errorless trials. Immediately following the conclusion of VD learning, a recall test was given in which the wrong syllable of each pair was presented individually while *S* attempted to recall the right syllable previously paired with it, in order to check on the extent to which intrapair associations had been learned during the VD task. No information was given as to the correctness of these recall responses, in order to prevent the actual learning of associations during the recall test.

PA procedure.—Having been given no previous indication that anything beyond the VD task would be required of him, *S* was taken immediately after the conclusion of VD learning to another part of the experimental room previously blocked from his view, and seated before the memory drum on which the PA task was presented. Standard instructions for PA anticipation learning were given, with an added statement that some of the syllables had been used previously on the VD task, so that *S* should try to anticipate the correct response the first time through the list. The PA learning task was identical for all *S*s, with each stimulus being presented alone for 2 sec., while *S* attempted to pronounce the correct response, followed by the stimulus-response pair for 2 sec., with an 8-sec. intertrial interval. Each *S* learned to a criterion of 2 successive errorless trials, or to a maximum of 24 trials.

Design.—A mixed-lists within-*S*s design was employed, whereby the 12 pairs were systematically varied in terms of their relationship to the previous VD task, as follows:

Discriminated (D) pairs: Four of the 12 PA pairs had also been used as pairs in the VD list. For 2 of these pairs, representing the D(S) condition, the PA stimulus had been "right" and the response "wrong" on the VD task. For the other 2 pairs, the D(R) condition, the response had been right and the stimulus wrong. As defined for purposes of this study, these pairs were assumed to have gone through both familiarization and discrimination phases during the VD task, but not the associative phase.

Familiarized (F) pairs: An additional four PA pairs consisted of syllables previously used in the VD list, but each had been paired therein with another syllable not used in the PA task. As with the D pairs, two of these F pairs consisted of stimuli which were right and responses wrong on the VD task, while the reverse was true for the other two pairs, which are designated as the F(S) and F(R) conditions, respectively. The familiarization phase was therefore assumed to have already been provided by the VD task for these pairs, but they had not been directly discriminated,

TABLE 1

VD LISTS ILLUSTRATING CONDITIONS FOR EXP. I AND II, AND PA LIST USED FOR ALL *Ss* OF BOTH EXPERIMENTS

Experiment I		PA List (Exp. I & II)		Experiment II	
VD List	Cond.			Cond.	VD List
R[a] W		S R[b]			R[a] W
BIY-ROJ	D(S)	BIY-ROJ		D(S)	BIY-ROJ
JAF-LIH	D(S)	JAF-LIH		D(S)	JAF-LIH
KUX-VOR	D(R)	VOR-KUX		D(R)	KUX-VOR
PIB-QAW	D(R)	QAW-PIB		D(R)	PIB-QAW
RUZ-ciq	F(S)	RUZ-TEY		L(S-R)	RUZ-ZOG
yun-TEY	F(S)	ZOG-DEQ		L	TEY-VAK
ZOG-qer	F(R)	PEH-VAK		L(S)	PEH-SUQ
gah-DEQ	F(R)	SUQ-GED		L(R)	GED-DEQ
bov-PEH	N	TIV-JEC		U(S-R)	TIV-QUM
VAK-zes	N	NUK-ZIL		U	JEC-YIS
jiz-SUQ	N	WOX-QUM		U(S)	WOX-ZIL
GED-jub	N	YIS-MUP		U(R)	MUP-NUK

Note.—VD syllables not used in PA list written in lower-case letters.
[a] R = right; W = wrong.
[b] S = stimulus; R = response.

either from the syllables paired with them or from any other syllable used in the PA list.

New (N) pairs: The final four PA pairs consisted of syllables which had not been used in the VD list, so that neither prior discrimination nor familiarization phases had taken place.

One of the six VD lists is presented in the left column of Table 1, to illustrate the method by which each condition was produced within the framework of a single PA list. By means of a latin square, each subset of four PA pairs was used under the D, F, and N conditions for one-third of the *Ss*. In addition, the right and wrong syllables of each VD pair were reversed for half of the *Ss* within each of these three subgroups, so that each PA pair also appeared equally often within the S and R subsets of the D and F conditions. As a consequence, a total of six different VD lists were used, each being learned by 6 *Ss*. Each of the first 6 *Ss* to appear in the laboratory was assigned to a different VD list on the basis of a predetermined random order, as was also true for each subsequent set of 6 *Ss*.

Experiment II

In order to check the possibility that the results of Exp. I were due to its within-*Ss* design, so that performance on the N condition of the PA task may have been facilitated by the relative novelty of these pairs, a separate control group of *Ss* was tested in Exp. II under conditions corresponding to the N condition of Exp. I. In addition, PA pairs consisting of syllables previously paired on the VD task with syllables from other pairs of the PA list were employed, replacing the F condition of Exp. I. Furthermore, besides the variation in the PA pairs of the stimulus or response position of the right and wrong syllables from the VD task, the number of right or wrong syllables within the PA pairs was also systematically varied. These and other minor differences from Exp. I are described more fully below, although in most respects the materials and procedure were identical for the two experiments.

Subjects.—The *Ss* were 60 summer session students who were paid for their services, and were assigned to an experimental group of 36 or a control group of 24 *Ss* in terms of a predetermined random order based on their order of appearance in the laboratory. An additional 8 *Ss* were eliminated due to their inability to learn the VD list to criterion within a reasonable time, and were replaced by the next *S* to be tested.

Materials.—The same nonsense syllables used in Exp. I were employed, along with an additional 16 syllables selected from the 32–41% range of association values (Archer, 1960) for the VD task of the control group.

Design.—Prior to PA learning, the control group learned a VD list consisting exclusively of syllables not used in the PA list. For the experimental group, all syllables of the PA task had also been used in the VD task, but the two tasks were related in several different ways. Four of the 12 PA pairs were identical to the D(S) and D(R) conditions of Exp. I. Four other PA pairs constituted a like-position (L) condition wherein each stimulus had been paired on the VD task with a stimulus from another PA pair, and each response with another response. Each of these pairs differed in the number or position of syllables which were right or wrong on the VD task, as follows: (*a*) L(S-R) condition: Both syllables of the PA pair had been right on the VD task; (*b*) L condition: Neither syllable was right on the VD task; (*c*) L(S) condition: The stimulus syllable was right and the response wrong; (*d*) L(R) condition: The response was right and the stimulus wrong. The final four pairs represented an unlike-position (U) condition, in that each PA stimulus had been paired on the VD task with a response from another PA pair, and each response with the stimulus from another pair.

As with the L conditions, these represented all possible combinations of right and wrong syllables, and were also subdivided into U(S-R), U, U(S), and U(R) pairs.

As in Exp. I, a latin square counterbalancing procedure was employed so that each pair of the PA list was used equally often under each condition. Due to the increase in number of conditions, the use of 12 different VD lists was necessary to accomplish this, each being learned by 3 *Ss* of the experimental group. One of these lists is presented in the right half of Table 1, to illustrate the method by which the various relationships between the VD and PA pairs were produced.

Other minor differences from Exp. I included: (*a*) the use of a different *E*; (*b*) an expansion of the recall test following the VD task so that both the right and wrong syllables of each pair were presented individually for recall; (*c*) an addition to the PA instructions of a statement that the PA task would be more difficult than the VD task, due to indications from Exp. I that *Ss* may have been unaware of the difference in difficulty, and that their resulting discouragement interfered with their PA performance.

Results

Since 83% of the *Ss* both in Exp. I and II failed to reach the criterion of 2 successive errorless trials on the PA task within 24 trials, total number of correct anticipation responses over the 24 trials represented the most appropriate PA performance measure. This was therefore computed separately for each of the 12 PA pairs for each *S*, assuming that those few *Ss* who reached criterion would have continued to respond correctly on all subsequent trials. A second measure, more appropriate as an index of initial rate of learning, was the number of errors (trials) preceding the first correct response to a given pair, designated as *before-errors*. In addition, the percentage of trials following the first correct response on which errors were made (*after-errors*) was also obtained. The means for each of these three measures are presented for the various conditions of Exp. I and II in Tables 2 and 3, respectively. Also examined was the nature of any overt errors on the PA task, which are not presented due to their infrequency and failure to show significant effects.

Although the use of different *S* populations precludes any direct comparison between the two experiments, their results will be discussed together, because of the close comparability between them in conditions and procedures. Unless otherwise stated, all comparisons reported as statistically significant refer to the .01 significance level, although all comparisons significant at the .05 level are reported as such.

D-F-N conditions (Exp. I).—Table 2 shows that performance for the N condition was numerically superior for all measures, with the exception of after-errors, where it was slightly surpassed by the D condition. Analyses of variance revealed significant differences between conditions for correct responses ($F = 8.34$) and after-errors ($F = 5.66$), but not for before-errors ($F = 1.75$, all Fs based on 2 and 150 *df*). Further analysis revealed that the D and N conditions did not differ significantly from each other, but both were significantly superior to the F condition in correct responses, as well as after-errors.

D-L-U conditions (Exp. II).—As shown by Table 3, performance under the D condition was superior on all measures to the L and U conditions, but significantly so only for correct responses ($F = 5.23, df = 2/264$). Neither before-errors ($F = 2.69$) nor after-errors ($F = 2.35$) differed significantly between conditions. Further analysis revealed that the D condition was significantly superior in correct responses to both the L and U conditions, which did not differ from each other.

Experimental-control group (Exp.

TABLE 2

PA Performance Measures for Conditions of Exp. I

Measure	D Condition			F Condition			N Condition
	D(S)	D(R)	Total	F(S)	F(R)	Total	
Correct responses	10.61	8.82	9.72	7.25	8.08	7.67	9.88
Before-errors	9.29	12.25	10.77	11.31	12.01	11.66	10.28
After-errors	.279	.249	.265	.429	.326	.378	.280

II).—The performance of the control group, corresponding to the N condition of Exp. I, was numerically superior to the L and U conditions on all measures, but slightly inferior to the D condition, as shown by Table 3. However, none of these differences approached significance by analysis of variance (all Fs <1). This lack of significance, despite the close comparability in magnitude of the differences involved with the previously described significant differences both in Exp. I and II, is a direct result of the use of different *S*s in the experimental and control groups, and the consequent necessary inclusion of between-*S*s variance in the error term appropriate for this analysis. Since the amount of total variance was quite similar within the experimental and control groups, the within-*S*s error variance from the previous D-L-U analysis involving the experimental group could be used as an estimate of error variance uncontaminated by differences between *S*s. When this was done, control group performance proved to be significantly superior to the combined L and U conditions ($F = 9.76$, $df = 1/264$), although differing at only the .05 level from the

TABLE 3

PA Performance Measures for Conditions of Exp. II

Measure	Control Group	D Condition		
		D(S)	D(R)	Total
Correct responses	10.57	11.22	10.63	10.92
Before-errors	9.91	9.37	9.88	9.63
After-errors	.250	.233	.248	.240

L Condition

Measure	L(S-R)	L	L(S)	L(R)	Total
Correct responses	9.92	8.31	9.00	9.83	9.26
Before-errors	9.69	12.14	10.22	10.14	10.55
After-errors	.307	.300	.347	.291	.312

U Condition

Measure	U(S-R)	U	U(S)	U(R)	Total
Correct responses	9.00	9.11	9.97	9.22	9.33
Before-errors	11.72	11.44	10.17	11.22	11.14
After-errors	.267	.274	.279	.279	.274

L ($F = 5.96$) and U ($F = 6.67$) conditions separately, and not at all from the D condition ($F < 1$). While the statistical legitimacy of this procedure is suspect, it at least serves to demonstrate that the results for the N (control) condition are closely comparable between Exp. I and II, and that the apparent discrepancies with respect to statistical significance are directly attributable to the greater efficiency of the within-*Ss* design of Exp. I.

Several points of agreement between Exp. I and II are noteworthy, particularly the fact that neither gave any indication of superiority for the D condition over the N (control group) condition. Moreover, all conditions wherein the syllables of the PA list were previously used in VD learning but paired differently (F, L, and U conditions) showed performance significantly below either the D or N (control) conditions. Finally, although the two experiments could not be compared directly, the difference from the D and N conditions was approximately equal for the F condition of Exp. I and for the L and U conditions of Exp. II, indicating that VD pairing with syllables from other PA pairs produces no more interference than does VD pairing with syllables not in the PA list.

VD right-wrong comparisons.—As shown in Table 2, the D(S) condition of Exp. I, wherein the stimulus of the PA pair was right on the VD task and the response wrong, was superior on all measures to the D(R) condition. However, the F(S) condition was superior to the F(R) condition only in before-errors. Analyses of variance showed that the combined D(S) and F(S) conditions were significantly superior at the .05 level to the D(R)-F(R) conditions in before-errors ($F = 5.99$, $df = 1/150$), but not for either correct responses ($F = 0.63$) or after-errors ($F = 1.98$). However, the D-F by (S)-(R) interaction proved to be significant at the .05 level for correct responses ($F = 4.74$, $df = 1/150$), which was shown by subsequent analysis to be due to the significant superiority of D(S) to D(R), while the inferiority of F(S) to F(R) was nonsignificant.

Table 3 shows that the D(S) condition was also numerically superior to D(R) for all measures of Exp. II, although not significantly so (all *F*s <1). Likewise, no consistent nor significant effect emerged from a comparison of the four possible combinations of stimulus and response, right or wrong within either the L or U conditions (these *F*s <1). With the possible exception of the L conditions, within which performance under the L(S-R) condition was best and the L condition worst, there is absolutely no evidence that PA performance is related to the number of syllables which were right on the VD task. Furthermore, in agreement with the F condition of Exp. I, PA learning was unaffected by whether the right syllable was in the stimulus or response position of the PA pair. Any such effect therefore appears to be limited to the D condition, although even here the superiority of the D(S) to the D(R) condition proved to be significant only in Exp. I.

VD recall test.—After VD learning, when *Ss* of Exp. I were presented the wrong syllable from each pair and attempted to recall the right syllable previously paired with it, an average of 17% correct associations were given, and only 7 of the 36 *Ss* failed to give any. In Exp. II, where this test was expanded to include the separate presentation of both the right and wrong syllable of each pair, one or more correct associations were made on 21% of the pairs for both the experimental

and control groups, while only 4 of the 36 experimental *S*s and 2 of the 24 control *S*s gave none. Each group gave correct associations to 13% of the 24 individual syllables of the VD task. Presentation of the wrong syllables, while *S* attempted to recall the right ones, produced 16% correct responses, as contrasted with 11% when the right syllables were presented. This difference was highly significant ($t = 5.24$, $df = 59$), indicating that *S*s had learned more about the right than the wrong syllables during VD learning, since they were better able to recall them.

Although relatively small in number, the extent to which correct intrapair associations were given on the VD recall test appeared to warrant a comparison in performance between PA pairs involving correct associations during VD recall (associated pairs), and those which did not (nonassociated pairs). Since such comparisons require performance measures both on PA pairs which did and did not involve correct VD associations, it was necessary to eliminate for this purpose those *S*s for whom all PA pairs within a given condition were either associated or nonassociated. For Exp. I, this left 18 *S*s within the D condition and 25 *S*s for the F condition, while the D, L, and U conditions of Exp. II included 26, 19, and 13 *S*s, respectively. It was found that more correct PA responses were given to associated than to nonassociated pairs under all conditions. For the D condition, this difference was significant at the .05 level for Exp. II ($t = 2.13$, $df = 25$) and for both experiments combined ($t = 2.35$, $df = 43$), but not for Exp. I ($t = 1.33$, $df = 17$). For the combined F, L, and U conditions, the superiority of performance on associated pairs was significant ($t = 2.78$, $df = 56$), even though these syllables had been paired differently on the VD task, so that different associations had to be formed during PA learning. However, this difference was not significant for either the F ($t = 1.93$, $df = 24$), L ($t = 1.35$, $df = 18$), or U condition ($t = 1.41$, $df = 12$) tested separately. No meaningful statistical comparison was possible between PA pairs with one and both syllables correctly associated on the VD recall test under the F, L, and U conditions, due to the small numbers involved, although slightly more correct responses were given to such pairs with both syllables associated correctly during VD recall than to corresponding pairs with one or no syllables correctly associated.

Correlation coefficients were computed between performance on the VD recall test and total correct responses on the PA task for *S*s of the various groups, to determine whether there was any relationship between them. A significant positive relationship was found for Exp. I ($r = .438$), and that for the experimental group of Exp. II was nearly significant ($r = .293$), but little or no relationship was shown for the Exp. II control group ($r = .104$). Apparently formation of associations on the VD task, as measured by performance on the VD recall test, affects subsequent PA performance only when the same syllables are used on both tasks, rather than representing a general tendency for fast PA learners to form more associations during VD learning.

VD performance.—One final attempt to explicate the relationships between PA and VD learning consisted of the calculation of the number of trials and errors to the criterion of two successive errorless trials on the VD task for each *S*, which were then correlated with the various PA performance measures, as well as the VD recall test. The VD performance measures were

closely comparable for Exp. I and the experimental and control groups of Exp. II. However, significant positive relationships were obtained between VD and PA performance as measured either by correct responses or before-errors only for Exp. I and the experimental group of Exp. II, with the various *r*'s ranging from .37 to .77. No significant relationships were found for the Exp. II control group, suggesting again that the use of similar material rather than a general learning ability accounts for the observed relationships between VD and PA performance. VD performance was also unrelated to scores on the VD recall test, indicating that increasing the number of presentations of the VD pairs did not affect the number of associations formed within these pairs.

Discussion

The present results provide no evidence for any facilitation of PA performance by prior VD learning involving the PA material. In terms of the original rationale underlying this study, it would therefore appear that familiarization and discrimination processes are unimportant in PA learning. Such a conclusion, however, is rendered unlikely by the numerous studies demonstrating positive transfer from a variety of familiarization and discrimination procedures prior to PA learning, although such findings are by no means universal (e.g., Arnoult, 1957; Underwood & Schulz, 1960). Nevertheless, if familiarization and discrimination processes do play a significant role in PA learning, it is clear that neither was produced to any substantial degree by the present VD task.

The absence of positive transfer from VD to PA learning in this and other studies[4] (Young & Underwood, 1954), along with the present finding of no relationship between VD and PA performance within the Exp. II control group, both indicate that basically different learning processes may be involved in the VD and PA tasks. Such an interpretation receives further support from the present comparisons involving the number and position of right and wrong VD syllables in the PA pairs, especially the finding that PA performance under the D(S) condition, for which the stimulus had been right and the repsonse wrong on the VD task, was significantly superior to the reversed D(R) condition in Exp. I. That more must be learned about the response than the stimulus in PA learning has been demonstrated previously (e.g., Underwood & Schulz, 1960). Moreover, the VD recall test of Exp. II clearly showed that more had been learned about the right than the wrong syllables during VD learning. Consequently, superior performance would be expected under the D(R) condition, where the right syllables, about which more had been learned, appeared as the PA response, about which more needs to be learned. The contrasting superiority of the D(S) condition suggests that whatever is learned about syllables on the VD task must be more appropriate to PA learning of the stimulus than to PA response learning, which also points to the probable locus of any differences between the VD and PA tasks in the basic learning processes involved.

That VD learning is in fact closely compatible with learning of the PA stimulus is indicated by the fundamental importance of recognizing and differentiating between items of the list in both cases. As concerns PA response learning, at least two previous investigators have suggested that this may be basically different from VD learning. Underwood et al. (1959) concluded that their method for producing PA response learning did not involve any substantial differentiation between responses, such as would be produced by the present VD task. In a similar vein, Runquist and Freeman (1960) have hypothesized that the process of connecting the component

[4] R. K. Young, personal communication, 1960.

letters of the response syllables to form a single response unit, termed "response integration" by Mandler (1954), may be largely absent from VD learning. Further evidence for the unimportance of response integration in VD learning is provided by the present finding of no relationship of VD performance to the results of the VD recall test, wherein response integration was necessary for successful performance. In contrast, response integration has previously been shown to constitute a major part of PA learning of unfamiliar material, such as the nonsense syllables employed herein (Underwood & Schulz, 1960), so that the principal difference between the VD and PA tasks may well lie in the extent to which this process is involved.

Differences with respect to response integration can also account for the obtained superiority of subsequent PA performance on pairs correctly associated during the VD recall test, inasmuch as successful VD recall implies that a substantial degree of response integration is present at the conclusion of VD learning, so that PA performance should be facilitated thereby. Rather than actually being produced as a result of VD learning, however, these correct VD recall associations may simply mean that the component syllables possessed greater familiarity or association value to the individual *S* prior to the experiment.

The significant negative transfer found under the F, L, and U conditions represents the most puzzling result of the present study, which defies obvious explanation. To the extent that intrapair associations had been learned on the VD task, interference of these VD associations with subsequent learning of the re-paired syllables on the PA task would offer a means of accounting for this finding, since negative transfer commonly occurs under such conditions (e.g., Porter & Duncan, 1953). However, the absence of any significant facilitating effects under the D condition, wherein the PA and VD pairs were identical, along with the relatively small number of intrapair associations shown by the VD recall test, together indicate that such associations could have been at best a minor factor in accounting for the observed negative effects on PA performance. Furthermore, performance under the combined F, L, and U conditions was actually significantly better on those PA pairs containing syllables correctly associated on the VD recall test, as compared with pairs involving no measurable learning of associations during the VD task. As suggested previously, this superior PA performance on pairs involving correctly associated VD syllables is quite likely due more to the greater prior familiarity or association value of these syllables than to the actual learning of associations during VD learning. Nevertheless, such a result is diametrically opposed to the formation of wrong associations on the VD task as an explanation for the inferiority of the F, L, and U conditions, and indicates that the primary locus of the observed interference must lie instead in the nonassociated VD syllables. Of course, the possibility remains that the VD recall test was not sufficiently sensitive to record weaker intrapair associations formed during VD learning, and the lack of any other apparent mechanism to account for the observed negative transfer effects lends credence to this suggestion. However, if the presence of negative transfer is used to infer that associations had been learned on the VD task, these associations were by definition of a partial or subthreshold variety, so that significant evidence would thereby be provided in opposition to the all-or-none theory of association formation (Estes, 1960; Rock, 1957).

Since the differences between the D condition and the combined F, L, and U conditions were as expected, but both were depressed with respect to the N (control) condition, an alternative possibility is that some general source of interference may have operated only on those PA pairs consisting of syllables used previously in the VD task. More specifically, the complex nature of the relationships of the various PA pairs to the VD task, introduced as a result of the present within-*S*s design, may

have served to confuse *S* and to interfere with his performance on these pairs as compared with the N pairs not used in the VD task. However, the virtual identity of the results of Exp. I and II, which differed primarily in this respect, argues against such an interpretation, as does a previous study wherein a within-*Ss* procedure strikingly similar to the present one yielded results indistinguishable from a between-groups design where each *S* served under only one condition (Twedt & Underwood, 1959).

Although the VD and PA tasks were made to differ considerably except for the communality of the nonsense-syllable material, in order to minimize the likelihood of contamination by nonspecific sources of transfer, any negative effects produced thereby should have affected equally the N (control) condition, and therefore cannot readily account for the depression of the other conditions relative to the N condition. Therefore, while an explanation in terms of interference on those syllables common to the VD and PA tasks would be entirely consistent with the present results, there is no direct evidence to support this position, and possible mechanisms which might produce such interference remain in the realm of speculation.

In conclusion, the present study has demonstrated that prior VD learning of nonsense-syllable pairs does not necessarily facilitate subsequent PA performance on those same pairs, and that PA performance may be significantly retarded when the syllables are paired differently on the VD and PA tasks. While no satisfactory explanation is evident for either finding, and the generality of these effects beyond the present conditions remains to be determined, it is apparent that basically different learning processes may be involved in VD and PA learning, and that the further examination and specification of these differences should represent a basic and important area for future investigation in verbal learning research.

Summary

In an attempt to determine the role of familiarization and discrimination processes in verbal paired-associate learning, two experiments were performed in which paired-associate (PA) learning of a nonsense-syllable list was preceded by verbal-discrimination (VD) learning, and the relationships between the syllables contained in the VD and PA pairs were systematically varied. Performance on PA pairs which were also used in the VD task (D condition) was significantly superior to conditions wherein the syllables of the PA pairs had been paired differently on the VD task, whether with other syllables not used in the PA list (F condition) or with syllables from other pairs of the PA list (L and U conditions). However, the D condition did not differ significantly from new pairs not used in the VD task with respect to PA performance, and this latter N (control) condition was significantly superior to the F, L, and U conditions. This latter negative transfer effect could not be attributed to the learning of "wrong" associations on the VD task, as measured by performance on a recall test at the conclusion of VD learning. In view of the absence of any positive effect of prior VD learning on subsequent PA performance, it was concluded that the two tasks may involve different learning processes, and some speculations were offered as to the possible nature of these processes.

References

Archer, E. J. A re-evaluation of the meaningfulness of all possible CVC trigrams. *Psychol. Monogr.*, 1960, **74**(10, Whole No. 497).

Arnoult, M. D. Stimulus predifferentiation: Some generalizations and hypotheses. *Psychol. Bull.*, 1957, **54**, 339–350.

Estes, W. K. Learning theory and the new "mental chemistry." *Psychol. Rev.*, 1960, **67**, 207–223.

Gibson, E. J. A systematic application of the concepts of generalization and differentiation to verbal learning. *Psychol. Rev.*, 1940, **47**, 196–229.

Hovland, C. I., & Kurtz, K. H. Experimental studies in rote-learning theory: X. Pre-learning syllable familiarization and the length-difficulty relationship. *J. exp. Psychol.*, 1952, **44**, 31–39.

Mandler, G. Response factors in human learning. *Psychol. Rev.*, 1954, **61**, 235–244.

PORTER, L. W., & DUNCAN, C. P. Negative transfer in verbal learning. *J. exp. Psychol.*, 1953, **46**, 61–64.

ROCK, I. The role of repetition in associative learning. *Amer. J. Psychol.*, 1957, **70**, 186–193.

RUNQUIST, W. N., & FREEMAN, M. Roles of associative value and syllable familiarization in verbal discrimination learning. *J. exp. Psychol.*, 1960, **59**, 396–401.

TWEDT, H. M., & UNDERWOOD, B. J. Mixed vs. unmixed lists in transfer studies. *J. exp. Psychol.*, 1959, **58**, 111–116.

UNDERWOOD, B. J., RUNQUIST, W. N., & SCHULZ, R. W. Response learning in paired-associate lists as a function of intralist similarity. *J. exp. Psychol.*, 1959, **58**, 70–78.

UNDERWOOD, B. J., & SCHULZ, R. W. Meaningfulness and verbal learning. Chicago: Lippincott, 1960.

YOUNG, R. K., & UNDERWOOD, B. J. Transfer in verbal materials with dissimilar stimuli and response similarity varied. *J. exp. Psychol.*, 1954, **47**, 153–159.

(Received December 12, 1960)

Journal of Experimental Psychology
1962, Vol. 63, No. 3, 269–274

RELATION BETWEEN INTERSESSION INTERVAL FREQUENCY OF COMPETING RESPONSES AND RATE OF LEARNING[1]

DAVID PREMACK AND RICHARD T. PUTNEY

University of Missouri

Although several studies have shown the rate of change in performance to vary inversely with the intersession interval (ISI) (e.g., Baron, 1952; Pennington & Thompson, 1958), more recent studies have shown the frequency of certain responses to vary directly with ISI (Premack & Bahwell, 1959; Premack & Collier, 1962). Responses that vary directly with ISI are those that, instead of showing an acquisition curve over sessions, tend to have their greatest frequency on the first test session, i.e., the so-called exploratory responses. Responses of this class quite probably represent the principal competing responses of learning situations. But if the exploratory responses of one situation are the competing responses of the learning situation, then the opposite effect of ISI on learning and exploratory-response frequency, respectively, suggests that the former may be accounted for by the latter. That is, the greater the ISI, the greater the frequency of competing responses, and therefore, the slower the rate of learning.

This hypothesis was tested in the present experiment by using the light-contingent bar press as a representative competing response in an acquisition situation in which rats were required to learn to drink. By placing *Ss* in the test space at different ISIs and equating water deprivation by interpolated home-cage drinking sessions, it was possible to determine the relation between bar press frequency and rate of acquisition of the drinking response.

METHOD

Subjects.—Thirty naive male albino rats of the Sprague-Dawley strain, CA 80–90 days old, were used as *Ss*. They were housed individually in wood-frame cages with hardware cloth backs, fronts, and floors.

Apparatus.—Each of 10 wooden boxes, built in a rack of two tiers of 5 each, was equipped with a drinking tube and light-contingent bar. Boxes were $7 \times 8\frac{1}{2} \times 8\frac{1}{2}$ in., painted flat black inside and out, and had hardware cloth floors. Illumination inside the box was virtually eliminated by fitting the refuse trays close to the base of each tier of boxes.

Drinking tubes were modified 50-ml. burettes, calibrated to 0.1 ml., mounted in the center of the hinged front door through which *S* was placed in the test box. The drinking spout had an inner diameter of 5 mm., and extended $\frac{3}{4}$ in. into the box, 2 in. above the floor. Like drinking tubes, similarly located, were used in home cage.

Each box contained an L shaped bar on the wall adjacent to the drinking tube. The bar protruded about 1 in. into the box, $2\frac{1}{2}$ in. above the floor, and was activated by about 10 gm. force. Each bar press turned on a 1-in. Dialco pilot light mounted in the ceiling directly above the bar, the light remaining on as long as the bar was depressed. Bar pressing was recorded by counters housed in a remote sound-treated box.

Procedure.—Pretraining: After a week of gentling, *Ss* were placed on an $11\frac{3}{4}$-hr. water deprivation schedule which was maintained throughout the experiment. Dry food was

[1] Part of the data reported here was taken from a Master of Arts thesis submitted by R. Putney to the Graduate School, University of Missouri. The study was begun during the senior author's tenure as a United States Public Health Service Postdoctoral Fellow and was completed with the support of Grant M-3345 from the National Institute of Mental Health.

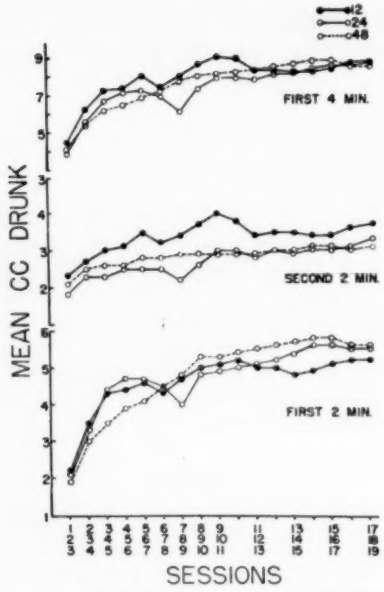

FIG. 1. Mean cc drunk averaged over running session triplets, for first 2 min., second 2 min., and first 4 min., with ISI as the parameter.

available in home cage at all times except during the drinking session, when food was removed. Drinking sessions, in both home and test cage, lasted 16 min. To insure that accommodation to the drinking schedule would not be confounded with subsequent test operations, 14 home drinking sessions were given on the deprivation schedule; these proved sufficient to stabilize home drinking. On Day 8, all *Ss* were given their first drinking session in the test box. The *Ss* were then equally divided into three matched groups with the matching variable being the mean proportion of the total for a session that was drunk in the first 2 min. averaged over the last four home sessions, plus the proportion of the total drunk in the first 6 min. of the first test session. The groups are referred to hereafter as 12, 24, and 48, the number denoting the number of hours that subsequently intervened between the successive test sessions which the groups received.

Testing: The experiment proper began on the morning of Day 9, with testing of the 12-hr. group. Thereafter the 12-hr. group was placed twice daily in the test boxes, at 8:30 A.M. and 8:30 P.M., and thus received no further home drinking. The 24-hr. group received water in test boxes every P.M., and in home cages every A.M., thereby maintaining the 11¾-hr. water deprivation schedule. The 48-hr. group received water in test boxes every other P.M., and had three home drinking sessions interpolated between every 2 test sessions. All groups received 20 test sessions, and after the first session each *S* was assigned to a test box in which it received all remaining sessions. Data recorded were: amount consumed in successive 2-min. intervals, both for home and test drinking; and number of bar presses by 2-min. intervals.

RESULTS

Acquisition of drinking.—Acquisition of the drinking response is shown in Fig. 1 where mean volume consumed is plotted against sessions with ISI as the parameter. Separate curves are shown for the first 2-min., second 2-min., and first 4-min. intervals, all of which have been smoothed by plotting the data quasicumulatively (see Fig. 1 for exact method). Examination of the data by successive 2-min. intervals showed that from the fourth minute on there were essentially no differences either between the groups or over sessions, and that the main consistent differences of both kinds were confined to the first 4 min. of a session. An analysis of variance of

TABLE 1

ANALYSIS OF VARIANCE FOR THE FIRST 4 MIN. OF TEST DRINKING OVER SESSIONS 1–7

Source	*df*	*MS*	*F*
Between *Ss*	29		
Interval (I)	2	13.38	2.42
Error (b)	27	5.53	
Within *Ss*	180		
Sessions (S)	6	161.62	118.66**
I × S	12	2.47	1.82*
Error (w)	162	1.36	
Total	209		

* $P = .05$.
** $P < .001$.

amount consumed in the first 4 min. for the first seven sessions is summarized in Table 1. The closeness of the means for the terminal sessions, 16 through 20, suggests that the groups stabilized at a common level; the significant interaction that was obtained between group and session indicates that the rate of approach to the terminal drinking level differed for the groups, being generally inversely proportional to the ISI, as shown in Fig. 1. When means for Sessions 16 through 20 are used as estimates of asymptotes, mean volumes consumed in the first 4 min. for Sessions 6 and 7 combined were: 93% of asymptote, 83%, and 77% for Groups 12, 24, and 48, respectively.

Although the groups differed in rate of approach to terminal drinking level with respect to *initial* volume consumed, there was no difference in their approach to the terminal drinking level in terms of *total* consumption. In an analysis of variance identical with the preceding one except for the use of total, rather than initial volume consumed, the Group × Session interaction failed to approach the 5% level ($F < 1.00$), and only the Session main effect was significant ($F = 20.86$, $df = 6/162$, $P < .001$). Thus the groups increased over test sessions both in initial and total volume consumed, but the rate at which they increased differed only for initial consumption.

Interpolated home-cage drinking.— Two additional analyses established that differences in test drinking between Groups 24 and 48 were *not* due to differences in home-cage drinking that might possibly have emerged during the interpolated home-cage drinking sessions which both of these groups received during the testing proper. An analysis of variance for group and session made on the first

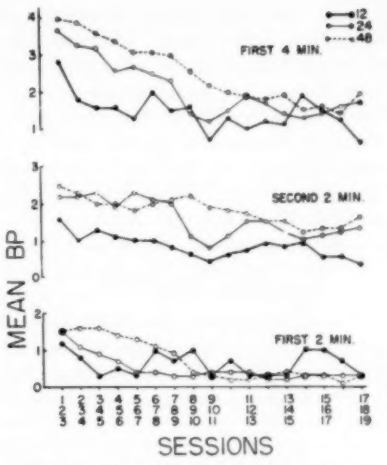

FIG. 2. Mean number of light-contingent bar presses averaged over running session triplets, for first 2 min., second 2 min., and first 4 min., with ISI as the parameter.

2 min. per session of the interpolated home-cage drinking for Groups 24 and 48 showed that neither the main effect of Group ($F = 1.96$, $df = 1/18$) nor the Group × Session interaction ($F = 1.61$, $df = 8/144$) attained the 5% level. Likewise, an identical analysis made on *total* amount drunk per session over the same period showed that neither the Group, nor the Group × Session interaction, attained the 5% level ($F < 1.00$, both cases). In sum, the analyses of drinking showed that the groups differed only in initial volume consumed in test drinking, and this in terms of rate of approach to the apparently common terminal drinking level. Neither total consumption per session in test drinking, nor initial or total consumption in interpolated home-cage drinking were found to differ.

Figure 2 shows mean bar press frequency plotted against sessions with ISI as the parameter; separate curves are shown for the first 2-min., second

2-min., and first 4-min. intervals, all of which have been smoothed by plotting the data quasicumulatively (see Fig. 2 for exact method). An analysis of variance, summarized in Table 2, was made on bar presses in the first 4 min. for ISI and Session over the same sessions as were used in the analysis of initial test drinking, viz., first seven sessions. Both Session and ISI main effects were significant, while their interaction failed to attain the 5% level, indicating that, although all groups declined over sessions and at about the same rate, the absolute frequency of bar pressing varied directly with the ISI.

An analysis identical with the preceding one was made on bar presses for the session rather than for the first 4 min.; the Session main effect proved to be significant ($F = 7.83$, $df = 6/162$, $P < .001$), indicating that all groups declined in total bar presses over the first seven sessions, but neither the main Group effect nor the Group × Session interaction approached the 5% level ($F < 1.00$, both cases). Final analyses were made for ISI on bar presses both in first 4 min. of the session and total session summed over the last seven sessions, the large number of zero responders in these terminal sessions making a test for trend inadvisable. No differences were found either for bar presses in first 4 min. or for the session ($F < 1.00$, both cases).

In sum, joint consideration of the drinking and bar press analyses bears out the hypothesis of an inverse relation between frequency of the light-contingent bar press and rate of approach to the terminal drinking level. (*a*) The groups differed in rate of increase in amount drunk in the first 4 min. over the first seven sessions and, at the same time, in frequency of the bar press, there being a generally inverse relation between mean initial amount drunk and mean initial bar press frequency; also, there was an over-all increase in initial drinking concurrent with an over-all decrease in initial bar pressing. (*b*) All groups increased in total amount drunk over the first seven sessions and at the same time, decreased in total frequency of bar pressing, and there were no reliable differences between groups either in amount consumed per session or in bar press frequency per session. (*c*) All groups appeared to approach approximately the same terminal drinking level both in initial and in total consumption, and correspondingly, all groups appeared to approach about the same terminal bar press level both in initial and total frequency.

TABLE 2

ANALYSIS OF VARIANCE FOR THE FIRST 4 MIN. OF BAR PRESSING OVER SESSIONS 1–7

Source	*df*	*MS*	*F*
Between *Ss*	29		
Interval (I)	2	62.96	3.78*
Error (b)	27	16.66	
Within *Ss*	180		
Sessions (S)	6	18.73	3.92**
I × S	12	3.41	.72
Error (w)	162	4.72	
Total	209		

* $P < .05$.
** $P < .001$.

DISCUSSION

Although the present results show a rather clear inverse relation between competing-response frequency and rate of change in the referent response, it may be asked whether ISI does not have additional effects upon learning independent of competing responses. Estes (1955), for example, would derive the present results for drinking, as well as a more general body of distributional learning phenomena, from a random stimulus fluctuation assumption in statistical learning theory. Pennington and Thompson (1958), who found an

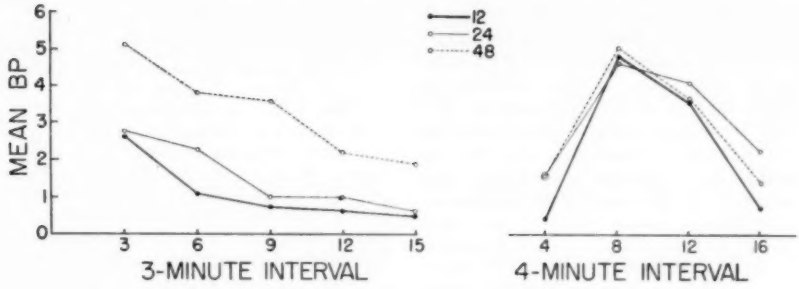

FIG. 3. Comparison of light-contingent bar pressing with and without the accompaniment of drinking. (See text for explanation.)

inverse relation between ISI and pattern discrimination learning, for ISIs identical with the present ones, ascribe the relation to a forgetting factor.

However, in none of the evidence cited for the noncompeting response interpretations of the ISI learning relation does there appear to have been any measurement of competing responses, and, of course, failure to measure competing responses does not establish that they are not present. Judged by the occasions on which competing responses have been observed in the learning situation (e.g., Cotton, 1953; Pereboom & Crawford, 1958), as well as the occasions on which their presence has been inferred to account for effects ranging from extinction to the partial reinforcement effect, it may be concluded that learning situations with, are decidedly more common than those without, competing responses. Further, though few learning situations contain the light-contingent bar press, its effect may be considered similar to that of the unmeasured competing responses that are typically present. But if it is reasonable to assume that competing responses are nearly ubiquitous, then, in order to consider whether they represent a principal, if not sole, determinant of the ISI learning relation, it need only be considered whether the positive relation between ISI and responses apt to enter the learning situation as competing responses is itself a general one. An answer to this question may be contained in the fact that in two species and three response systems, we have found no response, whose frequency is greatest on the first session, that does not vary directly with ISI (Premack & Bahwell, 1959; Premack & Collier, 1962).

The effect of the bar press on drinking is, however, only one side of the coin for there is, in addition, the effect of drinking on the bar press. Figure 3 presents a comparison of the bar press as it occurred in one case with, and in another without, the accompaniment of drinking. The right and left panels of Fig. 3 show mean light-contingent bar presses per 4-min. and 3-min. interval, respectively, averaged over the last 4 and last 6 sessions of series of 20 and 24 sessions, respectively; data in the left panel show the bar press without the accompaniment of drinking, and were taken from previous results (Premack & Collier, 1962). In addition to substantially delaying the peak bar press frequency, the presence of drinking ultimately overcame the ISI effect, so that in the terminal sessions, which these data represent, differences between Groups 12, 24, and 48 were negligible. The reciprocal effects between the bar press and drinking suggest that for at least some cases the two sides of response competition may be described by the following hypotheses. (*a*) Bar press on drinking: if the referent response has a higher independent rate (see Premack, 1959) than the competing response, its asymptote will be unaf-

fected but its rate of approach to asymptote will vary inversely with frequency of the competing response. (*b*) Drinking on bar press: if the referent response has a lower independent rate than the competing response, its asymptote will vary inversely, and its approach to asymptote directly, with frequency of the competing response.

Summary

The experiment was designed to test the hypothesis that the inverse relation between intersession interval and rate of learning may be due to a direct relation between ISI and frequency of those responses which seem most apt to enter the learning situation in the role of competing responses. In a space containing both a light-contingent bar and a drinking tube, three groups of water-deprived rats ($N = 10$) were tested repeatedly at ISIs of 12, 24, and 48 hr., respectively. Water deprivation was equated by interpolated home-cage drinking sessions, and measures were taken of both bar pressing and amount drunk. All groups approached about the same terminal drinking level, at a rate that was generally inversely proportional to their bar press frequency, which is in support of the hypothesis.

References

Baron, M. R. The effect of long inter-trial intervals on the limit of eyelid conditioning. *J. exp. Psychol.*, 1952, **44**, 438–441.

Cotton, J. W. Running time as a function of amount of food deprivation. *J. exp. Psychol.*, 1953, **46**, 188–198.

Estes, W. K. Statistical theory of distributional phenomena in learning. *Psychol. Rev.*, 1955, **62**, 369–377.

Pennington, D. F., & Thompson, R. Supplementary report: Discrimination learning in rats as a function of highly distributed trials. *J. exp. Psychol.*, 1958, **56**, 94–95.

Pereboom, A. C., & Crawford, B. M. Instrumental and competing behavior as a function of trials and reward magnitude. *J. exp. Psychol.*, 1958, **56**, 82–85.

Premack, D. Toward empirical behavior laws: I. Positive reinforcement. *Psychol. Rev.*, 1959, **66**, 219–233.

Premack, D., & Bahwell, R. Operant-level lever pressing by a monkey as a function of interest interval. *J. exp. Anal. Behav.*, 1959, **2**, 127–131.

Premack, D., & Collier, G. Joint effects of stimulus deprivation and intersession interval: Analysis of nonreinforcement variables affecting response probability, *Psychol., Monogr.*, 1962, **76**(5, Whole No. 524).

(Received December 14, 1960)

Journal of Experimental Psychology
1962, Vol. 63, No. 3, 275–277

SKIN CONDUCTANCE LEVELS AND VERBAL RECALL [1]

R. N. BERRY

Indiana University

In the present experiment the question is whether there are individual differences in skin conductance levels which are related to performance on a recall test. The skin conductance level may be taken to represent "activation" in at least one sector of the organism. There is evidence that "activation," as found by a number of physiological measures, is related to intraindividual variations in performance (Freeman, 1940; Stennett, 1957). Between *Ss*, however, the results could conceivably be different, since individual differences in performance measures may not be generated by differing somatic states, i.e., levels of skin conductance, and conductance, on the other hand, may be generated by factors unrelated to performance.

Malmo (1959, p. 376) has recently stated that: "It is important to stress that the measure denoted by 'moderate activation level' has meaning only in relative (not in absolute terms) terms. . . . Comparisons are invariably of the within-individual, within-task kind. . . ." This is a conservative approach, unquestionably motivated by experience and by technical considerations. However, Malmo (1958) in an earlier paper looked with favor on Schlosberg's (1954, p. 84) statement that: "Preoccupation with transient changes, the PGR, has led to general neglect of the slow drifts in absolute level of conductance, despite the fact that the absolute level is the obvious correlate of general level of tension or activation." It is not clear from Schlosberg's discussion whether he meant the statement to imply truly absolute value which would permit interindividual comparisons, or whether the intent was simply to place more emphasis on the slower changes in skin resistance.

Although the interindividual relation has not been established, neither has it been disproven, and the present experiments are a venture to test it. The type of relation found within individuals has usually been the inverted U curve showing maximum performance to occur with a moderate "activation" level. If there is an interindividual relation, one might expect it to be of this type.

METHOD

Subjects.—The 18 *Ss* of Exp. I were volunteer male students from the available summer session population. All except 1 or 2 *Ss* were paid on the basis of an hourly wage. The 32 *Ss* of Exp. II were also male volunteers, but they were unpaid and they were taken from the author's elementary class.

Apparatus and procedure.—The following description is appropriate for both experiments. Each *S* was seated in the experimental room in a straight-backed chair with a small table in front of him. The room was separated from the recording room by a sound-reducing partition. After *S* was seated and assured that he would not receive an electric shock, the recording electrodes were attached.

The electronic circuits for the apparatus, with some modifications, are those described by Davis, Siddons, and Stout (1954). One set of electrodes provided leads to an electrical impedance plethysmograph; the data from this are preliminary, have no obvious relation to skin conductance or to performance, and are not discussed in this report. The palmar skin resistance was continuously measured by means of another ink recording amplifying system. Nonpolarizing zinc, manganese

[1] This research was carried out under Contract Nonr 908(03). R. C. Davis is the principal investigator under this contract. A supplementary grant was provided by the Graduate School of Indiana University. The author is indebted to R. C. Davis, K. Kveim, and F. Berry for advice and material assistance.

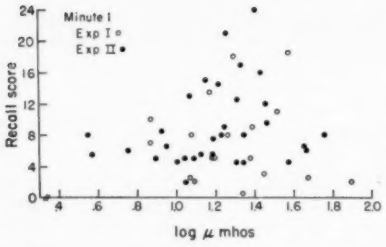

FIG. 1. The log conductance level of the first minute of the 10-min. learning session plotted against recall score.

dioxide, sodium chloride leads were placed on the left palm and upper left arm. Both skin areas were rubbed with alcohol before electrode placement, and the reference electrode on the upper arm was placed over a $7 \times 1\frac{1}{4}$ in. strip of $\frac{1}{2}$ in. thick foam rubber dampened with saline solution.

Following the electrode placement, *S* was given his instructions and the equipment operator made preparations to record. The *E*, in the experimental room with *S*, informed the equipment operator when the instructions were completed; the task session was then initiated.

The task.—The instructions to *S* were that he would be asked to memorize as many as he could of 30 paired associates. These were the final 30 paired items from a list by Melton and Safier, published by Hilgard (1951). Each pair had been typed on a 3×5 in. card and the deck of cards had been shuffled before the experiment was begun; each *S* was presented with this same order. The *S* was told, by means of sample cards that he would be required to recall the right-hand item of each of the 30 pairs when he was presented with the left-hand item. The latter was typed on another set of 3×5 in. cards and these had also been shuffled previous to the experiment; this shuffle determined the one order in which the cards were presented for the recall task.

The *S* was told that during the learning interval, he would have 10 sec. to study each pair and a 10-sec. intercard interval. He was further told that there would be a 6-min. interval between the learning and the recall session, and that he would be given 10 sec. to respond to each card in the recall period. All cards were presented manually on a small table in front of *S*.

Data analysis.—The midpoint values of each 1-min. interval from the beginning to the end of the task were read from the skin resistance records. These values were then converted to log conductance.

Results and Discussion

All of *S*'s verbal responses were recorded during the recall period and later divided into correct and incorrect categories. The only problem in deciding about the correctness of a response were the cases where the response was not exact but was obviously some grammatical form of the correct word; these responses were scored as half correct. With this scoring procedure the range of scores was from $\frac{1}{2}$ to 19 in Exp. I and 2 to 24 in Exp. II.

Following Exp. I, inspection of the data had indicated that simply dichotomizing the group into *S*s who scored better than 10 and those who scored 10 or less would be useful. All of the scores above 10 fell between the log conductance range of 1.1 and 1.6; this was defined as the optimum range. A 2×2 contingency table for Exp. I was set up using the above cutoff points, but Fisher's exact test showed a nonsignificant trend for the high scores to fall in the optimum conductance range.

The data of Exp. I looked sufficiently promising to justify a replication with more *S*s. The same cutoff

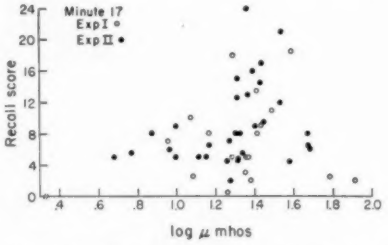

FIG. 2. The log conductance level of the first minute of the recall session (Min. 17 of the total experimental session) plotted against recall score.

points used in Exp. I were applied in a 2 × 2 contingency table for Exp. II and yielded an exact probability of observed frequencies of $P < .01$.

The log conductances for the midpoint of the first minute of the *recall* period are presented in Fig. 2. Using the same cutoffs as for Min. 1 the recall period conductance values are also predictive of recall scores; the Fisher exact test gave a $P < .05$ for the data of Exp. II. The P values for both Min. 1 and Min. 17 are one-tailed, but, since the direction of outcome was predicted from Exp. I, this is an appropriate procedure.

The important finding is that recall scores can be predicted from a measure of the organism's state in the first minute of learning. This probably represents an adjustment to the total experimental situation which is at least partially accomplished before the recording can be begun. That this adjustment persists is evident in the high correlation ($r_s = .80$ for Exp. I and .87 for Exp. II) between the conductance values of the first minute of learning and the first minute of recall.

The relation of skin resistance levels to recall scores clearly suggests the inverted **U** shape that has been previously found (Courts, 1942; Stennett, 1957) to relate level of activation (or arousal) to performance in complex tasks. However, it appears that optimum conductance levels do not, in many cases, insure a good performance. This latter finding must be cautiously interpreted because the resistance at the site of the reference electrode was not measured and could have varied considerably among *Ss*. Since this resistance enters into the total resistance recorded, it may be that the positive finding would have been even more substantial if the reference electrode resistance had been held constant.

Summary

In an effort to relate skin resistance to recall scores, two experiments, the second a replication of the first, presented male *Ss* with 30 pairs of meaningful words. There were 18 *Ss* in Exp. I and 32 in Exp. II. Each pair of words was presented once for 10 sec., with an interval of 10 sec. between pairs. Following this, there was a 6-min. delay before the reordered left-hand items were presented, 10 sec. at a time. The *S* was initially instructed to give as many of the right-hand items as he could remember, guessing if he wished. Skin resistance was continuously recorded throughout the experimental session and later converted to log conductance.

The data from Exp. I suggested that moderate levels of skin conductance in the first minute of the *learning* session were related to better recall. This relation was substantiated at a highly significant level ($P < .01$) in Exp. II. It was also found that moderate levels of conductance in the first minute of the *recall* period were optimal.

References

Courts, F. A. Relations between muscular tension and performance. *Psychol. Bull.*, 1942, **39,** 347–367.

Davis, R. C., Siddons, G. F., & Stout, G. L. Apparatus for recording autonomic states and changes. *Amer. J. Psychol.*, 1954, **67,** 343–352.

Freeman, G. L. The relationship between performance level and bodily activity level. *J. exp. Psychol.*, 1940, **26,** 602–608.

Hilgard, E. R. Methods and procedures in the study of learning. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951. Ch. 15.

Malmo, R. B. Measurement of drive: An unsolved problem in psychology. In *Nebraska Symposium on Motivation: 1958.* Lincoln: Univer. Nebraska Press, 1958.

Malmo, R. B. Activation: A neuropsychological dimension. *Psychol. Rev.*, 1959, **66,** 367–386.

Schlosberg, H. Three dimensions of emotion. *Psychol. Rev.*, 1954, **61,** 81–88.

Stennett, R. G. The relationship of performance level to level of arousal. *J. exp. Psychol.*, 1957, **54,** 54–61.

(Received December 15, 1960)

Journal of Experimental Psychology
1962, Vol. 63, No. 3, 278–283

THE EFFECT OF STIMULUS SIMILARITY ON THE ACQUISITION AND EXTINCTION OF A CONDITIONED RESPONSE [1]

DARWIN P. HUNT

Behavioral Sciences Laboratory, Wright-Patterson Air Force Base, Ohio

The research reported here was conducted in order to test two hypotheses: (*a*) the acquisition level of a conditioned response (CR) is inversely related to the similarity between the conditioned stimulus (CS) and a nonreinforced stimulus which is presented randomly during the acquisition or training phase and (*b*) the resistance to extinction of the CR is positively related to the similarity between the reinforced and the nonreinforced stimulus.

Consider the following situation: A 50% partial reinforcement schedule is arranged in a classical conditioning situation so that a tone of 1967 cps is always followed, 500 msec. after its onset, by a shock UCS; another tone of 1000 cps, randomly presented, is never followed by the UCS. Two other groups of *S*s are similarly conditioned, with the nonreinforced stimulus being a 1967-cps or a 153-cps tone. Following acquisition, all three groups are extinguished with the 1967-cps tone.

It might be assumed that during the acquisition phase two opposing tendencies are built up, one to respond and the other not to respond, to the 1967-cps tone. Each time a CS is reinforced, the tendency to respond to that CS (and other similar ones) is increased; each time a CS is not reinforced, the tendency not to respond to that CS (and other similar ones) is also increased. Although a number of curve forms have been proposed for the generalization gradient (Shepard, 1958), it will be assumed for predictive purposes that the tendency to respond (or not to respond) is a negatively accelerated decreasing function of the psychological dissimilarity of a stimulus to the reinforced (or nonreinforced) CS (Hovland, 1937).

Within this conceptual framework, the prediction concerning the relative differences among three such groups with regard to their acquisition level seems to be clear. Using the notation of Spence (1958), and ignoring drive, i.e., assuming it to be equal for all groups:

$$E = H - \bar{I}_n$$

E represents the excitatory potential associated with the reinforced stimulus.
H is the habit strength associated with the reinforced stimulus.
$\bar{I}_n$ represents the generalized inhibitory potential associated with the reinforced stimulus.

By assuming equal terminal levels of $\bar{I}_n$, i.e., the inhibitory tendency associated with the nonreinforced stimulus, $\bar{I}_n$ should decrease as the difference between the reinforced stimulus and nonreinforced stimulus increases. Then, if H is held constant, E would be predicted to be greater with larger differences between the reinforced stimulus and the nonreinforced stimulus.

With regard to the difference among the three groups in terms of the resistance to extinction of the CR to the 1967-cps tone, the prediction is not so straightforward. Humphreys (1939) found that resistance to extinction in a classical

[1] This paper is based upon a dissertation submitted to the Psychology Department of the Ohio State University. The author wishes to express his appreciation to D. D. Wickens, under whose guidance this research was conducted, for making available the GSR conditioning facilities and techniques developed under National Science Foundation Grant No. G-6203 and to W. D. Chiles for his comments and criticisms.

partial reinforcement situation was greater under a 50% reinforcement schedule than it was under a 100% reinforcement schedule. One could consider the 50% reinforcement situation as a discrimination learning task in which the nonreinforced stimulus is the same as the reinforced stimulus; while in the 100% reinforcement situation, the nonreinforced stimulus is represented by the "silence" or random stimulation occurring during the rest intervals. In this respect, then, these two situations represent the ends of a continuum of similarity of the reinforced stimulus to the nonreinforced stimulus, the 50% reinforcement situation representing the maximum similarity and the 100% reinforcement situation representing the minimum similarity between the reinforced stimulus and the nonreinforced stimulus.

Secondly, the apparent degree of change involved in switching from the acquisition to the extinction phase is greater for the *S* trained under a situation having a wide difference between the reinforced and the nonreinforced stimulus. Also, wide differences between the two stimuli would increase their discriminability during acquisition, which should serve to identify the stimulus which is followed by the reinforcement. Both of these factors, i.e., a greater apparent degree of change and the ease of discrimination, would be expected to operate in the direction of lowering the resistance to extinction. These two factors suggest that resistance to extinction, as related to stimulus similarity, should directly reflect the generalization gradient.

Finally, Wickens and Snide (1955) found that the nonreinforcement of an element (a light or a tone) of a reinforced stimulus complex (a light and a tone presented simultaneously) increased resistance of the CR to extinction. The above considerations are consistent in suggesting that resistance to extinction is a decreasing function of stimulus dissimilarity, i.e., the greater the difference between the reinforced and the nonreinforced stimulus, the less the resistance to extinction.

Method

Subjects.—A total of 125 men and women from a beginning course in psychology served as *S*s. Of these, 25 were eliminated for the following reasons: Procedural errors (14), *S*'s skin resistance rose beyond a previously determined level during the acquisition trials (6), the intensity of the shock was too great for *S*s to continue (3), and *S* failed to respond on all reinforced trials (2). The remaining 100 *S*s provided five groups, roughly equated with regard to sex, of 20 *S*s each. The rejected *S*s were distributed approximately equally over the five groups.

Apparatus.—The apparatus consisted of equipment for the presentation of a tone and a shock to *S* and for the measurement of the magnitude of the GSR accompanying these presentations.

Any one of three tones could be selected by *E* for presentation to *S* through earphones. The three tones differed physically both in frequency and intensity, but all had a duration of 600 msec. The frequencies were 153, 1000, and 1967 cps. The intensities of the three tones were adjusted so that they were equal in loudness to 1 sone, re: mean absolute threshold, as measured at the earphones. The rates of onset and decay of the tones were controlled by an electronic switch. The rise time was reduced until a click, which would result from a too rapid onset of the tone, was eliminated. Later measurement of this rise time revealed it to be about 35 msec.

Two shock electrodes were attached to *S*'s right hand, one to the tip of the forefinger and the other to the palm. The duration of the shock was 100 msec. The time interval between the onset of the tone and that of the shock was 500 msec.

Procedure.—The 100 *S*s were assigned to one of five groups on the basis of the order of their appearance for the experimental sessions and their pretest skin resistance level. Each *S* was given 10 reinforced trials during acquisition and then 10 extinction trials. On extinction trials, as well as nonreinforced acquisition trials, the shock was omitted.

There were three experimental groups, each of which had 10 nonreinforced trials during the acquisition phase in addition to the 10 reinforced trials. These nonreinforced trials were randomly interspersed among the reinforced trials, with the following restric-

tions: (*a*) No more than two reinforced or nonreinforced trials could occur consecutively, and (*b*) the first and last acquisition trials (Trials 1 and 20) were reinforced and Trial 19 was not reinforced. The pattern of reinforcement-nonreinforcement was the same for all three experimental groups. The primary difference among these three groups was the similarity of the nonreinforced stimulus to the reinforced CS, the 1967-cps tone. Group E-75 was given the 153-cps tone, another Group E-25 was given the 1000-cps tone, and Group E-0 was given the 1967-cps tone as the nonreinforced stimulus. These tones were selected because they are 75, 25, and 0 jnd's, respectively, removed from the reinforced CS (Hovland, 1937).

The other two groups served as control groups and were given only the 10 reinforced trials during the acquisition phase. In Group C-M the average intertrial intervals were equal to the average intertrial intervals of the three experimental groups. In Group C-S the intertrial intervals were equal to the intervals between reinforced trials of the experimental groups.

The intertrial intervals, measured between the onsets of successive presentations of the CS, were 25, 40, 55, or 70 sec. for all groups except Group C-S. These durations were randomly assigned to the 19 acquisition intertrial intervals of the three experimental groups and to the 9 extinction intertrial intervals. The interval between the final acquisition trial and the first extinction trial was 40 sec.

Upon the appearance of each *S*, a set of standardized instructions were read to him. After placing the GSR and shock electrodes on his hands, and the earphones over his ears, the 1967-cps tone was presented alone three times at 25-sec. intervals.

Immediately following the presentations of the tone, the voltage of the shock was adjusted according to the skin resistance level such that the amperage delivered to all *S*s was generally equated. The shock was then administered three times at 15-sec. intervals. In an attempt to compensate for the adaptation of *S* to the shock, the amperage was increased from approximately 1.5 to 2.5 ma. in 0.1-ma. increments over successive reinforced trials.

The base level skin resistance was determined to the nearest 200 ohms 70 sec. after the third administration of the pretest shock. According to his base level, each *S* was assigned to one of four classes: 0–28,000 ohms; 28,200–39,000 ohms; 39,200–50,800 ohms; and above 51,000 ohms. In the final design, each of the three experimental groups and each of the two control groups contained 5 *S*s in each class. Following this assignment, the acquisition and extinction trials were administered.

Response measures.—Two measures of response were employed. *Log conductance* (LC) values were computed as given by:

$$\log\left[\left(\frac{1}{R_a}-\frac{1}{R_b}\right)10^7+1\right]$$

R_b is the skin resistance (in ohms) immediately before the presentation of the CS.

R_a is the minimum skin resistance reached after the presentation of the CS and/or UCS.

Conditioning Index (CI) scores were obtained by the equation:

$$\frac{\text{LC}}{\frac{\sum_{t=4}^{10}\log\left[\left(\frac{1}{R_a}-\frac{1}{R_b}\right)10^7+1\right]}{7}}$$

where the numerator is the LC value for a particular trial and the denominator is the mean log conductance for a particular *S* for the last seven reinforced trials. Thus, a denominator was computed for each of the 100 *S*s. Analysis of the LC values on these seven reinforced trials for the three experimental groups indicated no difference among the groups.

Results

The means of the CIs during extinction as a function of trials are shown in Fig. 1. The differences in CIs among Groups E-0, E-25, and E-75 on the first extinction trial were evaluated by the Kruskal-Wallis one-way analysis of variance by ranks test (Siegel, 1956, p. 184) and found to be nonsignificant ($.20 > P > .10$).

The experimental design is such that three main variables are involved in the extinction phase: groups, base levels, and trials. Any particular *S* was included in only one group and one base level but was administered all 10 extinction trials. This design permits an analysis of the data by Lindquist's (1956, p. 281) Type III

analysis of variance design. The CIs during extinction were subjected to such an analysis. The statistical test that is crucial to the extinction hypothesis is the Group × Trial interaction. This interaction was significant ($P < .05$) whether all five groups or only the three experimental groups were included in the analysis, indicating that the rate of extinction was not the same for the different groups.

The fact that this interaction is significant and the observation that the relationship between the mean CIs and trials is approximately linear for all groups (Fig. 1) suggest that it would be appropriate and meaningful to fit a linear equation to the data for each group. A similar set of equations was computed using the LC values. The slope constants in these equations provide a quantitative estimate of the rate of extinction for each group.

The difference in linear trends among the three experimental groups based on the LC values was tested by computing a slope constant for each S in these groups and subjecting the resulting three arrays to Jonckheere's (1954) k-sample test. This yielded a z value of 2.17 which is significant ($P < .05$).

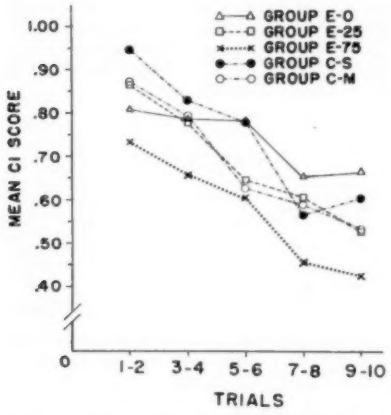

FIG. 1. The mean CI scores during extinction as a function of trials.

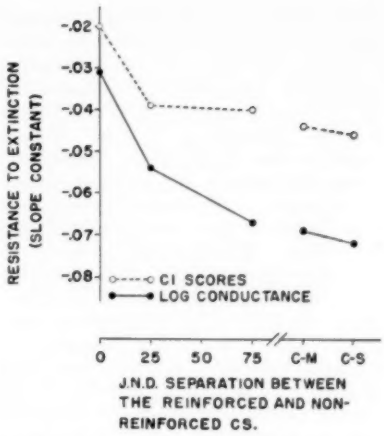

FIG. 2. Resistance to extinction as a function of the similarity between the nonreinforced and the reinforced stimulus.

These data can best be summarized by a plot showing the relationship between the slope constants and the stimulus similarity (Fig. 2).

DISCUSSION

In the initial portion of this paper two predictions were made; decreased similarity of the nonreinforced stimulus to the reinforced CS will result in (*a*) higher terminal acquisition levels of response to the previously reinforced CS and (*b*) lower resistance of this response to extinction. If the CI on the first extinction trial is an adequate index of the terminal acquisition level, and there is no apparent reason why this should not be the case, then support for the first prediction is lacking. Thus, a simple algebraic summation of these two tendencies, a positive response tendency associated with the reinforced stimulus and a negative response tendency associated with the nonreinforced stimulus, appears to be inadequate.

Other studies (Honig, Thomas, & Guttman, 1959) also have failed to obtain results supporting a simple summation hypothesis. This could be accounted for in several ways, in addition to the possible inadequacy of the Hullian theory in this situation. One possibility is that the number and/or ratio of reinforced and nonreinforced acquisition trials was such as not to permit discriminating responses to develop.

With regard to the prediction that decreased similarity between the reinforced CS and the nonreinforced stimulus should result in less resistance to extinction, positive support has been obtained. The significant ($P < .05$) Trial × Group interaction is consistent with this prediction. Using the slope constant, m, in the equation $y = mx + b$, as an index of resistance to extinction, i.e., the less negative the slope constant the greater the resistance to extinction, it can be seen in Fig. 2 that there is considerable agreement between the experimental findings and the original prediction. Using either the CI or LC values as the basis for the computation of the slope constants, results are obtained which are compatible with the original expectations. The slope constants based upon the mean LC values, however, produce a response gradient that more closely conforms to the original hypothesis than do those based upon the CIs.

The data suggest that the relative resistance to extinction of a CR to a "partially" reinforced CS can be predicted directly from knowledge of the generalization gradient involved. These results are compatible with the view that as the nonreinforced stimulus, in a discrimination learning task, becomes more similar to the reinforced stimulus, a classical partial reinforcement situation is increasingly approximated. The consequences of this are reflected in an increased resistance to extinction.

Summary

This study was designed to investigate the effects of the similarity between a reinforced and a nonreinforced stimulus on the acquisition and extinction of a conditioned GSR.

Five groups of 20 *Ss* each were given 10 reinforced acquisition trials consisting of a 1-sone, 1967-cps tone followed by a mild shock. Three of the groups were presented 10 nonreinforced trials randomly interspersed among the reinforced trials during acquisition. The treatment of these three experimental groups differed with respect to the similarity of the nonreinforced tone to the reinforced tone: 0, 25, or 75 jnd's separation. The other two groups were given no nonreinforced acquisition trials and differed from each other with respect to the duration of the intertrial intervals. Immediately following the acquisition phase, all *Ss* were given 10 extinction trials.

The basic response measures of changes in skin resistance were transformed into log conductance values and Conditioning Index (CI) scores for purposes of analysis. The transformation into CI scores tends to compensate statistically for the general "responsiveness" of the individual *Ss*.

No systematic relationship between the terminal acquisition level, as measured by the response on the first extinction trial, and stimulus similarity was found. However, a significant interaction between groups and trials during extinction was observed. The direction of this relationship was as predicted, i.e., the greater the similarity between the reinforced and nonreinforced stimulus the greater the resistance to extinction. It is suggested that this latter finding relates the concept of partial reinforcement to discrimination learning, i.e., as the nonreinforced stimulus in a discrimination task becomes more similar to the reinforced stimulus, a classical partial reinforcement situation is increasingly approximated.

References

Honig, W. K., Thomas, D. R., & Guttman, N. Differential effects of continuous extinction and discrimination training on the generalization gradient. *J. exp. Psychol.*, 1959, **58**, 145–152.

Hovland, C. I. The generalization of conditioned responses: I. The sensory generalization of conditioned responses with varying frequencies of tone. *J. gen. Psychol.*, 1937, **17**, 125–148.

Humphreys, L. G. The effect of random alternation of reinforcement on the acquisi-

tion and extinction of conditioned eyelid reactions. *J. exp. Psychol.*, 1939, **25,** 141–158.

JONCKHEERE, A. R. A distribution-free *k*-sample test against ordered alternatives. *Biometrika*, 1954, **41,** 133–145.

LINDQUIST, E. F. *Design and analysis of experiments in psychology and education.* Boston: Houghton Mifflin, 1956.

SHEPARD, W. N. Stimulus and response generalization: Deduction of the generalization gradient from a trace model. *Psychol. Rev.*, 1958, **65,** 242–256.

SIEGEL, S. *Nonparametric statistics for the behavioral sciences.* New York: McGraw-Hill, 1956.

SPENCE, K. W. Behavioral theory and selective learning. In M. R. Jones (Ed.), *Nebraska symposium on motivation: 1958.* Lincoln: Univer. Nebraska Press, 1958.

WICKENS, D. D., & SNIDE, J. D. The influence of non-reinforcement of a component of a complex stimulus on resistance to extinction of the complex itself. *J. exp. Psychol.*, 1955, **49,** 257–259.

(Received December 15, 1960)

Journal of Experimental Psychology
1962, Vol. 63, No. 3, 284–288

EVIDENCE FOR SHIFT EFFECTS IN THE CONSUMMATORY RESPONSE [1]

DAVID PREMACK AND W. A. HILLIX [2]

University of Missouri

The experiments reported here concern the effect of shifts in magnitude of sucrose concentration upon the mean rate of licking. The sucrose was presented directly, rather than being contingent upon a bar press, run, or choice, as in previous work (e.g., Collier & Marx, 1959; Crespi, 1942; Elliott, 1928; Young & Shuford, 1954; Zeaman, 1949); if the information sought is the effect of successive experiences with different stimuli, the tests suit the simplicity of the question to the extent that they avoid problems inherent in the instrumental response, and present the test stimuli as directly as possible. Further, the use of an instrumental response in previous work apparently diverted measurement to that response, for while previous work involved shifts in magnitudes of food or liquid, no measures appear to have been taken of *S*'s response to the food or liquid.[3]

METHOD

Subjects.—The *S*s were 36 hooded rats of both sexes, ca. 150 days old, from the University of Missouri colony. They were housed individually and maintained on ad lib. dry Purina chow and water, except for a daily 3-hr. period immediately preceding the test, when both food and water were removed from home cage.

Apparatus.—Apparatus consisted of a set of six adjoining boxes, each with a hole in the rear wall into which a drinking tube was inserted. Boxes were $7 \times 7 \times 9\frac{1}{2}$ in., built of wood except for hardware cloth floor and rear wall, and were painted flat black on all wood sides. Drinking tubes were modified 50-ml. burettes, with apertures of about 6 mm. inner diameter. A circuit connected to each tube was so arranged that each time *S*'s tongue contacted the fluid a minute current passed through *S*'s body, activating a counter. Counter readings were taken at 1-min. intervals.

Design.—Six groups were formed by factorially combining three concentrations of sucrose (4%, 16%, 32% by weight) and drinking sessions of two durations (2 and 15 min.). The *S*s were randomly assigned to the six groups with the restriction of equal *N* and like division of the sexes.

Procedure.—Testing consisted of two 20-day periods. In Period 1, each *S* received 20 daily sessions on either 4, 16, or 32% sucrose, a session lasting either 2 or 15 min. A shift to 32% was made on Day 21, so that in Period 2 all *S*s received 20 daily sessions of 32% sucrose, a session lasting either 2 or 15 min. as before. The *S*s were run in squads of six, with each of the six conditions represented in each squad, and with concentration and session-duration counterbalanced for apparatus. Following Period 2, two additional tests were run on the 15-min. groups, and these are described separately in the results section. Sucrose solutions were prepared on alternate days from commercial granulated sugar and distilled water.

[1] The research reported here was begun during the tenure of both D. Premack and W. A. Hillix as United States Public Health Service Postdoctoral Research Fellows, and was completed with the support of Grant M-3345 from the National Institute of Mental Health.

[2] Present address: Navy Electronics Laboratory, San Diego, California.

[3] A few studies measure lick rate as a function of sucrose concentration, but not as a function of changes or shifts in sucrose concentration. A study by Davis and Keehn (1959) treated momentary lick rate as a function of sucrose concentration on a within-*S* basis; no differences in momentary lick rate were found. Momentary consummatory rates in the rat appear to be largely constant (e.g., Stellar & Hill, 1952); in this article only mean consummatory rates are reported, and these may be considered to reflect mainly variation in either burst length or interburst interval.

RESULTS

Total licks per session.—In Fig. 1 mean lick rate per minute averaged over successive blocks of five sessions is plotted against blocks with sucrose concentration as the parameter; the upper and lower panels are for the groups receiving 2-min. and 15-min. sessions, respectively. Total licks made in the last five preshift sessions (Days 16–20) and first (Days 21–25) and last (Days 36–40) five postshift sessions were computed for each *S*, and analyses of variance were made on these data for concentration and period. Separate analyses were made for the 2-min. and 15-min. groups and the latter is summarized in Table 1.

For the 15-min., but not for the 2-min., groups both the main effect of Period and the Period × Concentration interaction were significant. Although Fig. 1 shows the ordinal relations of the postshift (but not preshift) results for the 2-min. groups to be essentially the same as those for the 15-min. groups, no significant

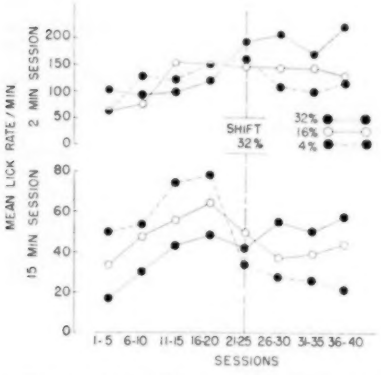

FIG. 1. Mean lick rate per minute averaged over successive blocks of five sessions with concentration (Days 1–20) and shift in concentration (Days 21–40) as parameters; upper and lower panels for groups receiving 2- and 15-min. sessions, respectively.

TABLE 1

ANALYSIS OF VARIANCE OF 15 MIN. GROUPS, FOR TOTAL LICKS PER SESSION

Source	*df*	*MS*	*F*
Between *Ss*	17	2,106,164	
Concentration (C)	2	777,961	.32
Error (b)	15	2,446,347	
Within *Ss*	36	2,563,861	
Period (P)	2	17,247,667	18.97*
C × P	4	7,860,051	8.64*
Error (w)	30	909,085	

* $P < .001$.

effects were found for the 2-min. groups, and the rest of this section concerns the 15-min. groups only.

In the preshift period (Days 16–20), when each group received either 4, 16, or 32%, there was an inverse relation between concentration and total licks per session, and differences between all pairs of concentrations were significant at the 5% level or better. Shifting the 4 and 16% groups to 32% produced a decrement in total licks per session, the decrement being greater for the 4% than for the 16% group. In the first postshift period (Days 21–25) there were no significant differences between any of the pairs of groups, but in the last postshift period (Days 36–40) only the difference between the original 16% and 32% groups failed to attain the 5% level. Thus the decrement in total licks per session produced in the 4 and 16% groups by a change to 32% was not transient, but rather, increased from the first 5 to the last 5 postshift days.

Lick rate in first 2 min.—Comparison of the 2-min. and 15-min. session groups was made by considering licks for the first 2 min. of the session only over the same periods as in the preceding analysis. An analysis of variance was made for session-duration (D) (2- vs. 15-min. groups), concentration (C), intrasession interval

(I) (1st and 2nd min.), and period (P). The only term involving D that proved to be significant was the P × C × D × I interaction ($F = 6.6$, $df = 4/10$, $P < .001$).

The significant over-all Period × Concentration interaction ($F = 4.9$, $df = 4/30$, $P < .005$), which is plotted in Fig. 2, confirms for first 2-min. lick rates what was earlier shown for session lick rates: the 32% control increased over periods while both shift groups declined; also, the curve for the 32% control makes it clear that, despite 20 preshift sessions, asymptotic licking had not been attained when the shift occurred.

Immediate shift effects.—Although the ultimate shift effects were decremental, immediate shift effects, occurring upon *S*'s first exposure to the new sucrose values, were incremental; the increments were obscured by the preceding analyses, concerned as they were with periods effects, but are shown by an analysis of lick rate in the first minute. A stabilized estimate of first-minute lick rate for the original concentrations was obtained by averaging over the last five preshift sessions (Days 16–20); these were compared with first-minute lick rate in the first (Session 21) and last (Day 40) postshift sessions, respectively. As is shown in Fig. 3, where first-minute lick rate is plotted for the above sessions with concentration as parameter, on the first shift-session there was a decided increase in both the original 4 and 16% groups, greater than the increase in the 32% control, and greater in the 4% than in the 16% group. The increment was transient, however, and, as is shown in Fig. 3, by Day 40 had become a decrement, both shift groups being then below the control. An analysis of variance made on these data for concentration, session duration, and period (as defined above), showed that both the main effect of Period ($F = 4.03$, $df = 2/60$, $P < .025$) and the Period × Concentration interaction ($F = 3.98$, $df = 4/60$, $P < .01$) were significant, confirming that the order of the groups changed over periods.

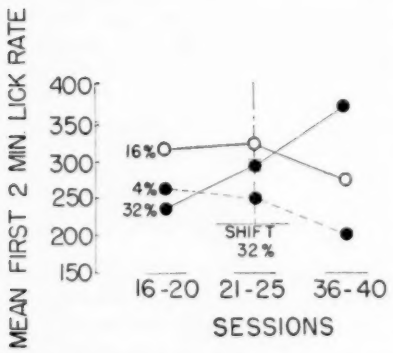

FIG. 2. Mean first 2-min. lick rate for designated blocks of five sessions with concentration (Days 16–20) and shift in concentration (Days 21–25 and 36–40) as parameters; 2- and 15-min. session groups combined.

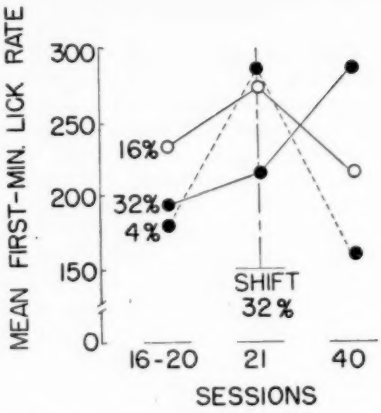

FIG. 3. Mean first-minute lick rate for designated sessions with concentration (Days 16–20) and shift in concentration (Days 21 and 40) as parameters; 2- and 15-min. session groups combined.

Permanence of shift effects.—To obtain further evidence as to the long-term stability of the shift effects, two further tests were run on the 15-min. session groups. Following Day 40, testing was discontinued for a period of 28 days. During this period, *Ss* were maintained ad lib., neither food nor water being removed from home cage in the daily 3-hr. period as had been the case. Following the 28-day rest, *Ss* were given 12 daily 15-min. sessions on the 32% solution as before. On the first postrest session, the groups did not differ in total licks per session, mean total licks for the original 4, 16, and 32% groups being: 546, 577, and 568, respectively. However, from the second session on the groups differed as before, mean total licks averaged over the 12 postrest sessions for the 4, 16, and 32% groups being: 381, 550, and 686, respectively. An analysis of variance on these data for total licks per session summed over the 12 sessions showed that the groups differed significantly ($F = 6.77$, $df = 2/15$, $P < .01$).

A second test was performed upon the 15-min. session groups to determine the effect of a change in food maintenance. All *Ss* were changed from the previous near ad lib. maintenance to a 23-hr. food deprivation schedule. After 3 days of the feeding schedule during which testing was suspended, *Ss* were given six daily 15-min. sessions on the 32% as before, the feeding schedule being maintained over the 6 test days. From the first session the groups differed as before. Mean total licks for the 4, 16, and 32% groups for the first session and averaged over the six sessions were: 729, 963, and 1,140, first session, and 893, 1,020, and 1,416, all six sessions. An analysis of variance on these data for total licks per session summed over the six sessions showed that the groups differed significantly ($F = 8.92$, $df = 2/15$, $P < .01$). Thus the follow-up tests showed that the decremental shift effect was (*a*) recoverable after 28 days of no exposure to the test stimulus, and (*b*) maintained after transfer from near ad lib. maintenance to a 23-hr. food deprivation schedule.

Discussion

The results serve to establish that consummatory responding is itself subject to shift phenomena, that the mean rate at which sucrose of a given concentration is licked depends in part upon previously experienced stimulus values.

The traditional generalization concerning shift effects would appear to be that their direction, particularly, and proportional magnitude to a lesser extent, are predictable from the stimulus scale as such. That is, changes in a positive direction on the stimulus scale are thought to produce positive response changes—overshooting—and negative stimulus changes to produce negative response changes—undershooting. However, the present data do not accord with this expectation for a positive stimulus change resulted in a negative response change: groups shifted from 4 and 16% to 32% performed *below* the 32 % control.

The impression that shift effects have a directional correspondence with the stimulus scale may be the result of a generally positive correlation between stimulus values and response probabilities and, in fact, not the stimulus value, but the response probability, may be the actual basis for prediction. If so, this could be determined only in cases where the stimulus and response values were not in perfect positive correlation. Such a case is provided by the present sucrose data for, as shown in Fig. 1, there was a perfect inverse relation between sucrose concentration and mean licks for the session. In the light of this, it is of interest that the sucrose data can be described as follows: shifts from stimuli associated with higher mean response frequencies to those of lower mean response frequencies resulted in negative response changes. Though, of course, it is conjectural as to whether shift effects have a directional correspondence with a scale of response probabilities, it is nevertheless evident that traditional assumptions concerning the stimulus scale will not mediate the present sucrose results.

Summary

Mean lick rate was studied as a function of changes in sucrose concentration. Rats given repeated 15-min. free drinking sessions on 4 and 16% sucrose, when shifted to 32% sucrose, showed immediate but transient incremental shift effects in first-minute lick rate, followed by long-term decremental shift

effects in both first-minute and session lick rates. The decremental effects did not abate over 20 daily sessions, and were not removed either by absenting *Ss* from the test stimulus for 28 days or by a change in food maintenance.

References

Collier, G., & Marx, M. H. Changes in performance as a function of shifts in the magnitude of reinforcement. *J. exp. Psychol.*, 1959, **57,** 305–309.

Crespi, L. P. Quantitative variation of incentive and performance in the white rat. *Amer. J. Psychol.*, 1942, **55,** 467–517.

Davis, J. D., & Keehn, D. Magnitude of reinforcement and consummatory behavior. *Science*, 1959, **130,** 269–271.

Elliott, M. H. The effect of change of reward on the maze performance of rats. *U. Calif. Publ. Psychol.*, 1928, **4,** 19–30.

Stellar, E., & Hill, J. H. The rat's rate of drinking as a function of water deprivation. *J. comp. physiol. Psychol.*, 1952, **45,** 96–102.

Young, P. T., & Schuford, E. H. Intensity, duration, and repetition of hedonic processes as related to acquisition of motives. *J. comp. physiol. Psychol.*, 1954, **47,** 298–305.

Zeaman, D. Response latency as a function of the amount of reinforcement. *J. exp. Psychol.*, 1949, **39,** 466–483.

(Received December 15, 1960)

Journal of Experimental Psychology
1962, Vol. 63, No. 3, 289–293

REACTION TIME AS A MEASURE OF INTERSENSORY FACILITATION [1]

MAURICE HERSHENSON [2]

Brooklyn College

In measuring reaction time (RT) to simultaneously presented stimuli, Todd (1912) found facilitation when sound and electric shock were paired but no effect when light was paired with either sound or shock. These results may be understood using the scheme suggested by Hilgard (1933) to account for facilitation of the eyeblink reflex. Hilgard used light and sound stimuli slightly offset in time and found that the magnitude of the eyeblink to sound was augmented when the light preceded the sound by intervals of 25 and 50 msec. Thus, he maintained, there was a summation of "reaction tendencies" when the stimuli were presented at an interval equal to the differences in the latencies of their respective responses. Since the latency for light is greater than that for sound, the light must precede the sound in order for "response tendencies" to occur simultaneously. Indeed, Hilgard suggests that the difference between the latencies for the single stimuli may be the point of maximum facilitation.

When taken together these two studies present a coherent picture. Since sound and shock yielded very similar RTs, no stimulus onset asynchrony (Δt) was necessary for simultaneity of response tendencies to result in facilitation. Similarly, facilitation should not have been expected for light and sound or light and shock when presented simultaneously since light yielded much longer RTs than either sound or shock.

The experiment to be reported was designed to study the effects of stimulus onset asynchrony on RT to paired light and sound stimuli. According to the above analysis, facilitation should be expected when the light and sound stimuli are offset in time by an amount approximately equal to the difference between the RTs to the single stimuli. The RT to this presentation should be faster than to either stimulus alone and the maximum effect should occur when Δt equals the difference between the RTs to the stimuli when presented singly. Since stimulus intensity is an important determiner of RT to single stimuli, variation of intensity might also effect facilitation. This was another variable used in the study.

METHOD

Light.—The light stimulus was produced by pulsing a glow modulator tube (Sylvania R-1131C) in the plate circuit of a 6L6 vacuum tube with rectangular pulses from a Tektronix pulse generator. The glow modulator tube was mounted in a light-tight box with a circular aperture, 1 cm. in diameter, directly in front of the tube. The light was diffused over a ground glass plate placed in front of the lamp and pressed against the inside front face of the box. The aperture, viewed binocularly from a headrest 50 cm. away, subtended 1° 10′ of arc.

[1] This study was based on a thesis presented to the Department of Psychology of Brooklyn College, in partial fulfillment of the requirements for the master's degree. The author is indebted to Elizabeth Fehrer and David H. Raab for their guidance and assistance throughout the course of this investigation. This research was supported, in part, by grants from the National Institute of Neurological Diseases and Blindness (B-1028) and the National Science Foundation (G-6456).

[2] Now at Yale University.

TABLE 1

MEAN RTs (MSEC.) FOR SUBJECT SR

Δt (Msec.)	Stimulus						Light & Sound Minus Δt
	Light		Light & Sound		Sound		
	Mean	*SD*	Mean	*SD*	Mean	*SD*	
0	149	11.1	98	6.3	97	6.8	98
5	144	10.6	99	5.3	94	5.5	94
15	144	8.3	111	6.3	94	7.8	96
25	144	8.5	121	6.2	94	6.4	96
35	149	7.8	131	5.9	96	7.3	96
45	149	10.1	138	5.2	97	8.0	93
55	147	8.3	141	6.6	100	5.8	86
65	153	10.5	144	4.7	104	7.0	79
75	147	8.2	142	4.2	97	8.1	67
85	145	7.8	144	8.2	99	7.1	59

Note.—Stimuli: L & S; $\bar{L} - \bar{S} = 49.8$ msec. Each cell mean was based on 72 RTs.

Two luminances were used covering two log units. The unfiltered stimulus (L) was approximately 3000 ft-L as measured by a Macbeth illuminometer at the ground glass plate. An ND 2 Kodak filter was inserted between the glow modulator tube and the ground glass screen to produce the weaker stimulus (l). Light flashes were all 250 msec. in duration. A white cardboard ring, .75 cm. wide and 6.5 cm. in outside diameter, surrounded the stimulus aperture and served as a fixation device.

Sound.—A noise burst of 50-msec. duration was presented by closing a high speed relay with pulses from another Tektronix pulse generator. This stimulus, produced by a Grason-Stadler noise generator (Model 455B), was delivered through an audio amplifier to a set of Telephonics headphones. Two sound levels were used for both *S*s. The sensation levels corresponding to these intensities were 95 db. (S) and 75 db. (s) for Subject SR, and 85 db. (S) and 65 db. (s) for Subject MH.

Timing.—Both pulse generators were triggered from a Tektronix wave-form generator which allowed for the introduction of stimulus onset asynchrony. All critical timing functions depended on voltages which were regulated to better than 0.1%. All durations and time intervals were checked by means of a Berkeley counter (Model 7350). The RTs were measured by the counter to ±1 msec. For Subject MH, the counter was started by the light pulse for all measurements. For Subject SR, the counter was started in a similar manner except for the sound RT where the sound pulse triggered the counter. In all cases, the counter was stopped when *S* depressed a normally closed telegraph key.

Procedure.—The *S* was seated in a soundproof and lightproof room which was dimly illuminated from above and behind *S* by a 6-v. lamp enclosed in a light-tight box with a 5.5-cm. circular opening covered with several sheets of tissue paper. The luminance at the paper surface measured 22 ft-L. This lamp was 120 cm. from the ground glass plate.

Each trial was introduced by a 1000-cycle warning tone lasting 1 sec. which was delivered through a loud speaker in the chamber. This was followed by one of three silent foreperiods of 2, 2¼, or 2½ sec. selected according to a previously recorded random order. The stimulus was presented at the end of this foreperiod. The entire sequence was controlled by Hunter decade interval timers connected to present one trial every 13 sec.

A block of trials consisted of 20 reactions to a particular stimulus. Three blocks—light alone, sound alone, and the combination of light and sound with a particular preset asynchrony—made up a series. Two series were given in each session which lasted for approximately 1 hr. The sequence of the blocks was changed from series to series.

Each session began with 5 min. of dark adaptation. Four practice trials were given before each block and a 2-min. rest period was given between blocks. The *S* was instructed to respond as quickly as possible to the first stimulus perceived. Introspective

reports indicated that for all asynchronies the stimuli seemed "simultaneous."

The longest and shortest RTs for each block were discarded leaving 18 RTs per block. Each asynchrony was tested over two sessions (i.e., four blocks per condition) which yielded a total of 216 RTs (72 light, 72 sound, 72 light-sound).

Subjects.—The *S*s were one undergraduate student (SR) and the author. Both were well practiced.

Results

Table 1 contains mean RTs and *SD*s for Subject SR. The values of asynchrony indicate the amount by which the light preceded the sound for the pair RTs. In no instance did the sound precede the light.

The small differences in the mean RTs to light alone and sound alone reflect day-to-day variations. RTs to the light-sound pairs were measured from the onset of the light. If these means represented responses to the sound alone (no facilitation) we would expect the RT to increase by the amount that the light preceded the sound (Δt). The values should get progressively larger therefore, until the point where the RT to the combination equals the RT to light. The means should then remain fairly constant.

Values for RTs to the light-sound pairs minus the asynchrony indicate response latencies as if measured from the onset of the sound. Evidence for facilitation can be seen at those asynchronies where the RT to light-sound measured from the onset of the light is faster than to light alone and RT to light-sound measured from the onset of the sound is faster than to sound alone.

Figure 1 shows the curves of facili-

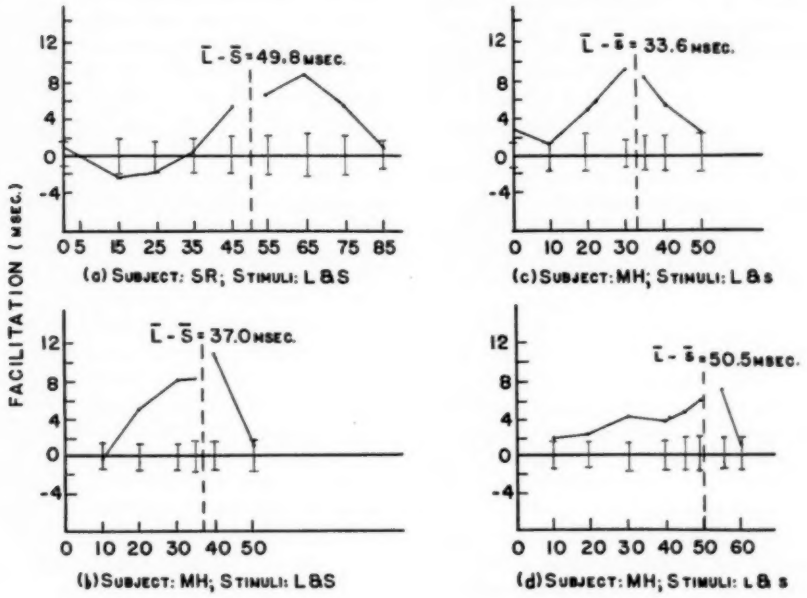
FACILITATION (MSEC.)
12
8
4
0
-4
L̄ - S̄ = 49.8 MSEC.
0 5 15 25 35 45 55 65 75 85
(a) SUBJECT: SR; STIMULI: L & S
L̄ - S̄ = 33.6 MSEC.
0 10 20 30 40 50
(c) SUBJECT: MH; STIMULI: L & S
L̄ - S̄ = 37.0 MSEC.
0 10 20 30 40 50
(b) SUBJECT: MH; STIMULI: L & S
L̄ - S̄ = 50.5 MSEC.
0 10 20 30 40 50 60
(d) SUBJECT: MH; STIMULI: L & S
STIMULUS ONSET ASYNCHRONY (MSEC.)

FIG. 1. Facilitation as a function of stimulus onset asynchrony.

tation as a function of asynchrony. These were derived from Table 1 and similar tables for Subject MH in the following way: first, overall means for all sessions were obtained for the RTs to light alone and sound alone (the best estimate of the RTs). The difference between these means is represented on the graphs as a vertical dashed line. Comparisons to the right of this line were made to light alone and those to the left of the line were made to sound alone. All comparisons were made using mean RTs which were obtained during the same sessions thereby minimizing the effects of practice and day-to-day variability.

Ordinate values were obtained simply by subtracting the RT to the pair from that to the single stimulus in accordance with the scheme outlined above. Thus a zero difference indicates no effect, a positive value indicates facilitation, and a negative value indicates inhibition. Fiducial intervals ($P = .05$) were calculated separately for each single RT mean with which comparisons were made at each asynchrony. The extent of each interval is indicated above and below the zero facilitation baseline.

It is evident that RT to paired light and sound stimuli can be facilitated by the introduction of stimulus onset asynchrony. All four curves indicate that the facilitation was a U shaped function of stimulus asynchrony with maximum facilitation in the region of the point at which the asynchrony is equal to the difference between the RTs to the single stimuli. These results are consistent with our hypotheses.

Individual differences are easily noted by comparing the curves for the most intense stimuli for 2 different *S*s (Fig. 1, a and b). The fact that the RTs to the same single stimuli were different for the 2 *S*s is reflected in the shift of the dashed line. A more important but less understandable shift was observed for Subject SR who showed maximum facilitation a full 15 msec. after the point of "physiological simultaneity." Thus for the same physical stimuli, RTs for the 2 *S*s were differentially affected by stimulus asynchrony whereas the magnitude of the facilitation was quite similar.

A comparison of Curves b and c in Fig. 1 (both curves for Subject MH) indicates that a change of 2 log units in the intensity of the sound had only a minor influence on the magnitude of the facilitation. However, the curves shifted appropriately in relation to the difference between the mean light RT and the mean sound RT, i.e., as expected, the maxima always fell near this point. On the other hand, examination of Curves c and d in Fig. 1 (also for Subject MH) reveals that a diminution of 2 log units in the intensity of the light stimulus yielded less facilitation. An appropriate shift in the maximum effect is again evident.

The *SD*s were quite small throughout indicating fairly consistent reactions to all stimuli (see Table 1). RTs to sound were generally more consistent than those to light. The *SD*s of the pair RTs were generally smaller than those of the RTs to the single stimuli at those asynchronies where facilitation was found. This trend fell short of significance however, a *t* test for correlated means comparing the pair *SD*s to the smaller of the two single stimulus *SD*s in the region of facilitation yielded a one-tailed *P* of .07.

Discussion

The data clearly demonstrate intersensory facilitation for paired suprathreshold light and sound stimuli. It

is further evident that this effect will be manifest only when one "gives the stimuli a chance," i.e., introduces the proper amount of stimulus onset asynchrony Since the amount of asynchrony necessary for facilitation is determined by the difference between the means of the RTs to the stimuli when presented singly, the value will differ from *S* to *S*.

That facilitation was not due to attitudinal factors is indicated by the fact that it was a function of stimulus asynchrony and that all four curves were similar in shape. If the mere presence of one stimulus had an "alerting" or "orienting" effect on the reaction to the other stimulus, one would have expected this effect to show up at all asynchronies that were measured. Although the data cover only a small range of asynchrony, it is obvious that this was not the case.

The trend in the *SD*s may be further evidence that facilitation was not due to an alerting function of one of the two stimuli for, if this were so, it is hard to see how *SD* should be related to asynchrony. It should be kept in mind however, that this trend was not statistically significant and must await further study before it can be incorporated into our conception of facilitation.

Another finding that is difficult to understand is the different effect of intensity of the two stimuli. Why should a 2 log unit change in the luminance of the light effect the magnitude of facilitation whereas a similar change in the sound showed no effect?

There is some hint of inhibition in the data. The negative values (Fig. 1a) show that the RTs to the combination were slower than to either stimulus alone. Although this effect seems to be rather small, the values at 25 and 50 msec. asynchrony are statistically significant and indicate that this region warrants further investigation.

Summary

An experiment is reported in which RTs were measured to light and sound presented singly and in combination. The chief experimental variable was stimulus onset asynchrony. The results may be summarized as follows: (*a*) Intersensory facilitation was demonstrated in that, for certain asynchronies, RTs were faster to the combination than to either stimulus alone. (*b*) Maximum facilitation occurred at or just beyond the point at which the asynchrony was equal to the difference in RT to the single stimuli. (*c*) Varying the intensity of the stimuli differentially affected the magnitude of facilitation. Reducing the light intensity resulted in less facilitation whereas reducing the sound intensity had no effect.

References

Hilgard, E. R. Reinforcement and inhibition of eyelid reflexes. *J. gen. Psychol.*, 1933, **8**, 85–113.

Todd, J. W. Reaction to multiple stimuli. *Arch. Psychol., N. Y.*, 1912, **3**, No. 25.

(Received December 16, 1960)

Journal of Experimental Psychology
1962, Vol. 63, No. 3, 294–301

VISUAL PERCEPTION OF THE HORIZONTAL DURING PROLONGED EXPOSURE TO RADIAL ACCELERATION ON A CENTRIFUGE [1]

BRANT CLARK AND ASHTON GRAYBIEL

San Jose State College

United States Naval School of Aviation Medicine

Correlates of the visual perception of the horizontal and vertical have been studied in considerable detail by numerous investigators for well over 100 years (Clark, Graybiel, & MacCorquodale, 1948; Graybiel, 1952; Noble, 1949). One device which has been used in this field in recent years is the human centrifuge. In certain experiments *S* is seated on the centrifuge some distance from the center of rotation, in a fixed position with respect to the gravitational vertical, and facing the direction of rotation. If he is then subjected to a gradual increase in the velocity of rotation, he feels that he is being pressed into the seat and that the seat is being tilted away from the center of rotation. Furthermore, if *S* observes a luminous, horizontal line in the dark under these conditions, the line will appear to be displaced in the direction of the apparent tilt (Clark & Graybiel, 1951; Noble, 1949; Witkin, 1944). The amount of this apparent displacement is related directly to the deviation of the resultant force from the gravitational vertical although orientation to the resultant force does not keep pace with the physical events. This particular reorientation of the visual horizontal is a special case of the oculogravic illusion and has been observed in flight as well as on a human centrifuge (Clark & Graybiel, 1949; Clark et al., 1948).

Although this phenomenon has been known since the time of Mach (Graybiel, 1952; Noble, 1949), only two studies have been made of the oculogravic illusion during prolonged exposure to centrifugal force (Clark & Graybiel, 1951; Graybiel & Brown, 1951). In these experiments the longest period of constant rotation was 15 min., and systematic changes in the illusion related to time of exposure to change in resultant force were not found. Similar results were found for the postural vertical during flight in an aircraft by MacCorquodale (1948). In related experiments using a tilting chair, the results of experiments on exposure to tilt from the vertical are conflicting. On the one hand, Passey and Guedry (1949) found a negative aftereffect, while Pearson and Hauty (1960) reported that delay up to 60 sec. was not a factor in *S*'s judgment of the postural vertical.

It was the purpose of this investigation to study the visual perception of the horizontal on a centrifuge during periods of change in the magnitude and direction of resultant force acting on *S* for a period of 4 hr.

METHOD

Subjects.—Four experienced *S*s including the authors and two volunteers were tested. The latter were naval enlisted men assigned to the United States Naval School of Aviation

[1] From the United States Naval School of Aviation Medicine, United States Naval Aviation Medical Center, Pensacola, Florida. Opinions and conclusions contained in this paper are those of the authors and do not necessarily reflect the views or endorsement of the Navy Department.

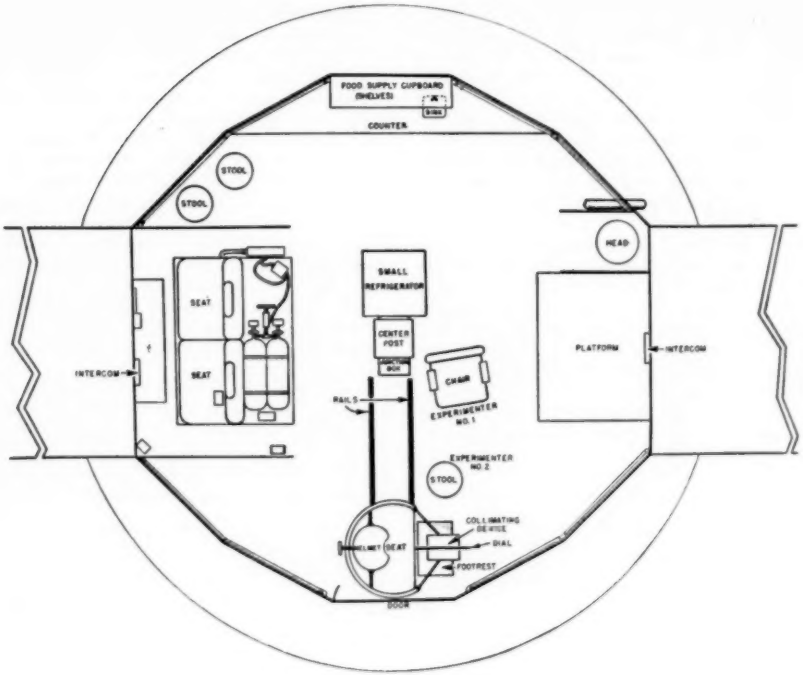

FIG. 1. Plan of the room.

Medicine with primary duty to act as *S*s in research projects. All 4 *S*s were thoroughly familiar with the problem under study and had made many other observations during previous rotations.

Apparatus.—All of the data were collected in the Pensacola slow rotation room (Graybiel, Clark, & Zarriello, 1960). This room, which is windowless and nearly circular, is 15 ft. in diameter and 7 ft. high. It is constructed around the center post of a human centrifuge, provides space for laboratory equipment, and includes living accommodations. Thus, it was possible for *E*s to be with *S*, and there was no limitation on the duration of the experimental period. The *E*s sat in positions indicated in Fig. 1, but they could move about the room at will. The *S* was seated facing in the direction of rotation in a specially constructed seat mounted firmly so that the center of *S*'s head was 6.25 ft. from the center of rotation (Fig. 1). The rotation was always counterclockwise, with deviations from the desired velocity being less than 2.5%. The seat consisted of three basic parts: (*a*) a Fiberglas head holder for each *S* molded to fit the head closely, and a chin strap to hold it rigidly in position; (*b*) a bucket seat and cloth support which could be adjusted to support each *S*'s back, the seat being adjustable vertically by a hand pump to press *S*'s head firmly into the head holder; and (*c*) a collimated luminous line of light (Clark & Graybiel, 1951) mounted on the chair directly in front of *S*. This line of light could be rotated in either direction about its center by *E* or by *S*. A luminous dial indicated the position of the luminous line and could be read by *E* to .25°. The apparatus made it possible for *S* to indicate his judgment of the horizontal by setting a line of light to horizontal in the dark during rotation, while being held rigidly in position to maintain constant throughout the testing period the relation of the head to the resultant force.

Procedure.—Each series of observations began with the centrifuge stationary. The *S* set the luminous line to horizontal in the dark 10 times, alternating clockwise and counterclockwise adjustments to establish

his point of subjective horizontal (PSH) under static conditions. This served as the point of reference from which all subsequent measurements were made. The PSH was established by determining the mean of the deviations. The room lights were turned on between trials to make readings in one series of observations, and the room was completely dark throughout the experimental period in others. The room was accelerated slowly to establish 6° or 12° deviation of resultant force (ϕ) from the gravitational vertical. The times were approximately 58 sec. for $\phi = 6°$, and 69 sec. for $\phi = 12°$. This force was the vector sum of the gravitational and centrifugal forces. Time for the run was counted from the time the room attained the required velocity (approximately 7 or 10 rpm) which was maintained for 4 hr. Beginning 1 min. after the constant velocity was attained, at 10 min., and at every 10 min. thereafter during the 4-hr. period of constant rotation, *S* made a series of four settings of the line to horizontal, two clockwise and two counterclockwise. The mean of these four settings was considered to be his new PSH. Measurements of the PSH were delayed for 1 min. to permit the effects of stimulation of the semicircular canals to be reduced to a a negligible amount. At the end of 4 hr. the room was stopped slowly (in approximately 38 sec. for $\phi = 6°$, and 49 sec. for $\phi = 12°$), and time was counted from this point. Four settings to the horizontal began at 1 min., and then were obtained at 10 min., 20 min., and 30 min. Between trials, *S*s smoked, listened to recordings, or talked to the 2 *E*s. Each *S* made one run with $\phi = 6°$ and a second with $\phi = 12°$, the lights being on between trials, and a third with $\phi = 12°$, the room lights being off during the whole observation period.

This procedure resulted in measurements of *S*'s perception of horizontal (the extent of the oculogravic illusion) in the absence of a visual framework under three conditions: (*a*) with the normal gravitational force, (*b*) during a prolonged period of constant increase

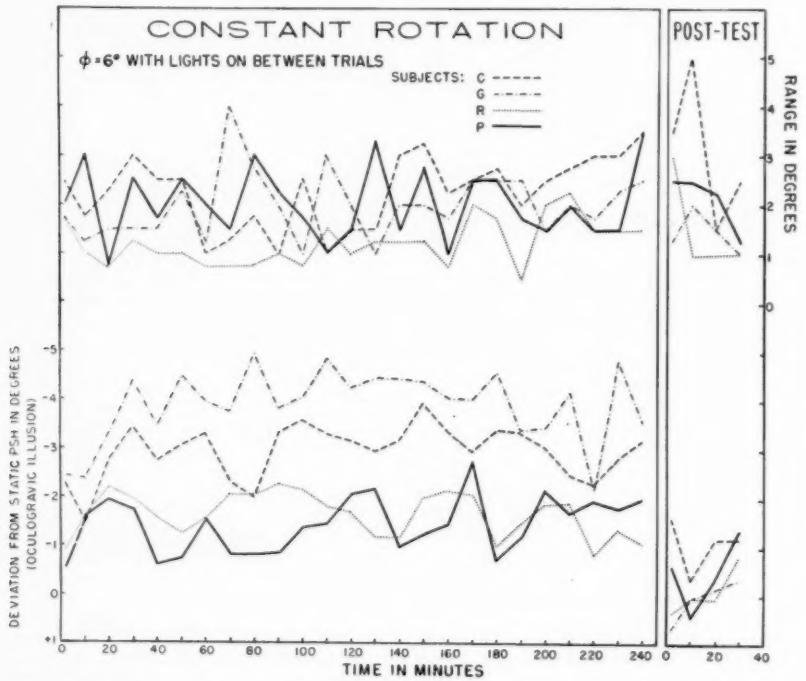

FIG. 2. The point of subjective horizontal (PSH) during and following 4 hr. of constant rotation, $\phi = 6°$, lights on between trials.

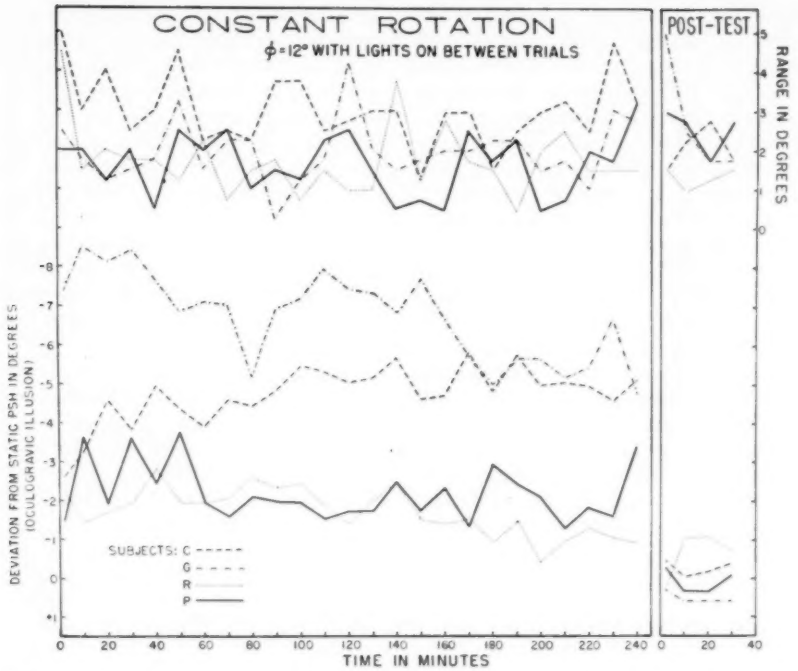

FIG. 3. The point of subjective horizontal (PSH) during and following 4 hr. of constant rotation, $\phi = 12°$, lights on between trials.

in centripetal force, and (*c*) with normal gravitational conditions immediately following cessation of rotation. Comparisons could then be made between these three conditions and the PSH (determined statically) and the two procedures used to make the measurements. In analyzing the data, evaluation of the effects of adaptation was based on time of stimulation by a constant centripetal force.

Each *S*'s point of subjective horizontal (PSH) was computed by calculating the mean of the 10 static settings of the luminous line to horizontal. This PSH served as the point from which all subsequent measurements in the series were made and is therefore the zero point of all of the graphs presented. As the rotation began, *S*s perceived the line to be displaced clockwise (considered to be plus in direction) to a new position which would remain relatively constant; i.e., *S*s observed the oculogravic illusion. In order to measure the extent of this deviation, *S* rotated the line counterclockwise (considered to be minus) to the apparent horizontal, and the mean deviation of the four settings from the static PSH was considered to be a measure of the oculogravic illusion for each point in time. Computations of median deviations also were made, but since there were few widely deviant scores, these data differed very little from the mean deviations. Therefore, plots of the mean values against time of the 4 *S*s and the experimental conditions constitute the basic data for the experiment.

RESULTS

Adaptation.—The primary purpose of this experiment was to determine whether a prolonged change in the magnitude and direction of the resultant force acting on human *S*s on a centrifuge would cause any change in the oculogravic illusion or in *S*'s PSH. If adaptive processes were involved, it would be expected

that the oculogravic illusion would tend to decrease systematically with time of stimulation. On the other hand, if the oculogravic illusion remained constant throughout a 4-hr. series of observations, it would be strong evidence that adaptation did not occur within these time limits.

The data (Fig. 2–5) show clearly that there were no systematic changes in the PSH with time; i.e., there appear to be no adaptative effects for the oculogravic illusion. The group data for $\phi = 6°$ exhibit no systematic change with time whereas one set of observations at $\phi = 12°$ shows a slight increase in the oculogravic illusion while the other shows a decrease. An examination of the data for individual *S*s (Fig. 2–4) also supports the notion that adaptation does not occur. A particular *S* may show a reduction in the illusion on one series of observations, but in another condition an increase or no systematic change may occur. An analysis of the data by computing the mean deviation for the six measures of the PSH for each of the four 1-hr. periods for each *S* was made, and the results further support the analysis of the graphic data. In toto, the data indicate clearly that adaptive processes did not occur in this particular perceptual situation within a 4-hr. period of constant stimulation. Furthermore there appear to be no aftereffects of this change in the resultant force since the observations for 30 min. following the period of constant rotation showed minimal deviations from the initial static PSH (Fig. 2–5).

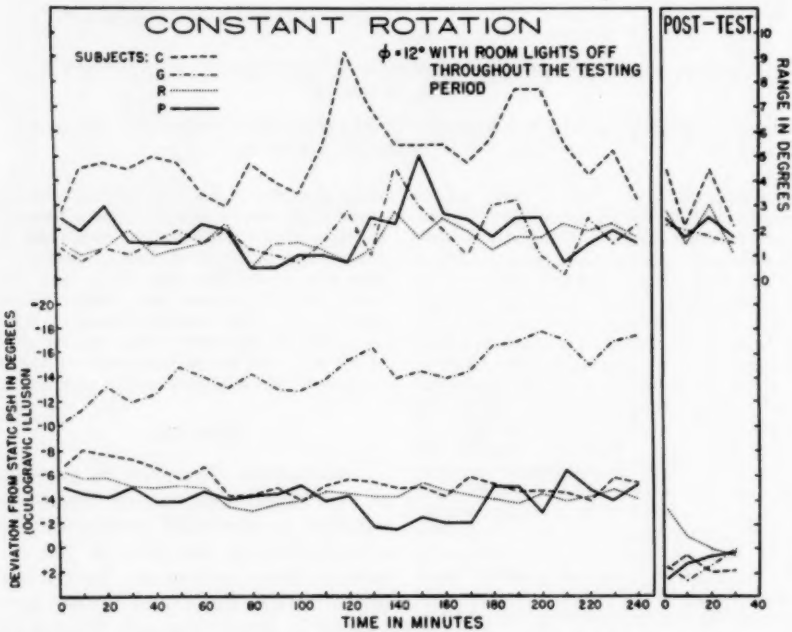

FIG. 4. The point of subjective horizontal (PSH) during and following 4 hr. of constant rotation, $\phi = 12°$, continuous darkness.

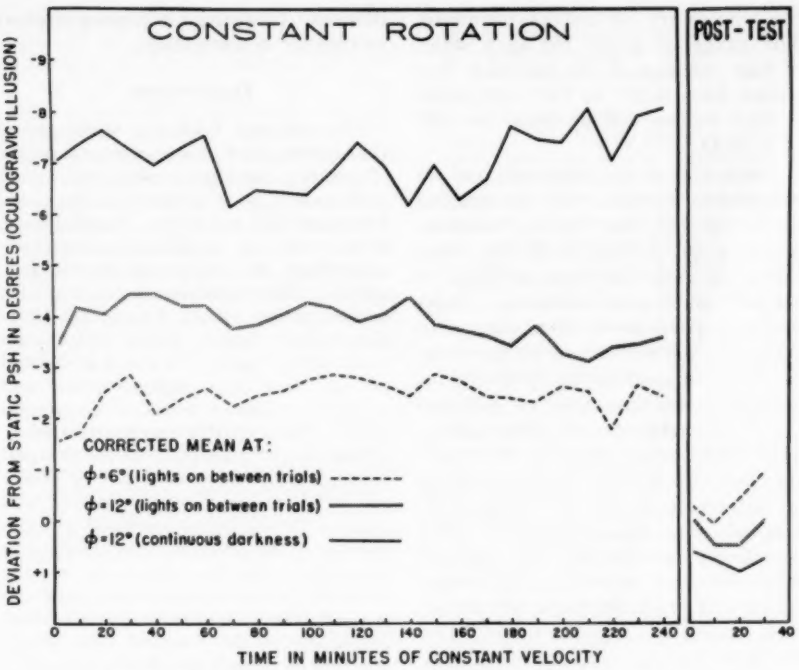

FIG. 5. The point of subjective horizontal (PSH) for 4 *S*s during the three stimulus conditions.

Relation of the oculogravic illusion to angle ϕ.—These results confirm the findings of earlier studies (Clark & Graybiel, 1951; Graybiel & Brown, 1951) which have shown that the oculogravic illusion is related to the angle ϕ, the angle between the direction of resultant force and the direction of gravity. An analysis of the group curves (Fig. 5) shows that the illusion at $\phi = 12°$ is consistently greater than at $\phi = 6°$. Similarly, the mean value for each of the 4 *S*s is greater at $\phi = 12°$ (Table 1). The results also show that the illusion measured in degrees is substantially less than the angle ϕ. The means of the 24 trials for each of the 4 *S*s beginning at 10 min. of constant rotation are given in Table 1. It will be noted that all of these values are substantially less than ϕ with the exception of 1 *S* when $\phi = 12°$ (Fig. 4). Plots of the range of scores for each series of four settings were also made to get some indication of

TABLE 1

MEAN DEVIATION IN DEGREES OF THE SETTINGS FROM THE STATIC PSH

S	Lights on between Trials		Continual Dark
	$\phi = 6°$	$\phi = 12°$	$\phi = 12°$
C	−2.96	−4.82	−5.38
G	−3.95	−6.70	−14.63
P	−1.85	−2.23	−3.96
R	−1.65	−1.69	−4.37
Mean	−2.60	−3.86	−7.09

the variability of *Ss*' performance. The range of scores for each series of four settings of the luminous line varied from 0.25° to 9.0° with most of the values falling close to 2.0° (Fig. 2–4).

Antecedent visual framework and the oculogravic illusion.—In the original design of the experiment, measurements were planned with the room lights on only between settings of the luminous line to horizontal. However, comparisons of these data with preliminary observations on the same *Ss* made in continuous darkness indicated gross discrepancies between the two methods of observation. Therefore, it was decided to repeat the runs with $\phi = 12°$ with *Ss* making all observations on the same time schedule in continual darkness. This completely eliminated the regular reference to the visual horizontal which was available to *Ss* just before each setting in the other series. The visual framework consisted of a grey wall with a horizontal aluminum bar around the room at about eye-level. The floor, ceiling, and vertical studs also were visible, but the visual field was partly obscured by the collimating device which was directly in front of *S*.

Analysis of the oculogravic illusion for the two sets of conditions for $\phi = 12°$ demonstrates that the oculogravic illusion was consistently greater when the room was continuously dark (Fig. 3–5). There is no overlapping in the two curves for the group data nor is there overlapping for *Ss* G, P, and R. The fourth *S* showed overlapping of two conditions (Fig. 3 and 4), but the mean value for the final 24 measures of the PSH is greater during darkness. These data show quite clearly that a prior visual framework before a setting has a marked influence on the oculogravic illusion. This variable is being studied in another investigation.

Discussion

The evidence from this study shows that within the 4-hr. experimental period of constant centripetal force, adaptation with respect to *Ss*' estimate of the visual horizontal did not occur. Furthermore, there were no significant, systematic aftereffects of this experimental procedure. These results are consistent with earlier findings (Clark & Graybiel, 1951; Graybiel & Brown, 1951) which were made with *S* isolated in a seat at the end of the centrifuge superstructure and exposed to higher levels of centripetal force. These results also are related to studies of the perception of the postural vertical using short intervals of bodily tilt. An early study by Passey and Guedry (1949) found increased error in settings to the postural vertical following prolonged delay in returning to the vertical. A more recent study by Pearson and Hauty (1960) has suggested that judgments of the postural vertical are due less to sensory adaptation than to a number of other factors. The data of the present study show that the visual perception of the horizontal under the influence of centripetal force is influenced by preceding visual stimulation; i.e., the oculogravic illusion increases with a reduction in prior visual framework. This indicates that if this estimation of the visual horizontal is regarded as a special case of the oculogravic illusion, the illusion is also a function of a number of complex factors other than the input from the otolith organs.

Differences among *Ss*' points of subjective horizontal for the three experimental conditions were conspicuous and consistent. The *Ss* could be ranked readily not only in terms of the magnitude of the oculogravic illusion but also in terms of the individual variability and influence of prior visual stimulation. This individual consistency suggests that the 4 *Ss* used different frames of reference in making the judgment. The frame of reference whether dependent on

visualization of the horizontal, on otolith cues, or on proprioceptive stimulation appears to be fairly consistent for each *S*, although individual *S*s showed what seem to be temporary changes in their frames of reference by occasionally making a consistent, temporary shift in the degree of the illusion in a succession of settings. The data from this study do not present evidence with regard to the specific character of the variables which produce these individual differences, but the evidence appears clear that the oculogravic illusion is a result of a fairly complex perceptual process.

Summary

Although there have been many studies on perception of the horizontal and vertical, no data are available on the effects of prolonged exposure to centripetal acceleration. In the present study of visual perception of the horizontal, 4 *S*s seated on a specially constructed seat in the Pensacola slow rotation room were studied during periods of constant rotation lasting for 4 hr. They were subjected to a constant change in magnitude and direction of centripetal force so that they perceived the oculogravic illusion.

The results show no significant, systematic changes in the oculogravic illusion during the period of constant rotation or during a 30-min. period following rotation. It was concluded that adaptation does not occur under these experimental conditions and that the oculogravic illusion is dependent upon a number of perceptual processes.

References

Clark, B., & Graybiel, A. Apparent rotation of a fixed target associated with linear acceleration in flight. *Amer. J. Ophthal.*, 1949, **32,** 549–557.

Clark, B., & Graybiel, A. Visual perception of the horizontal following exposure to radial acceleration on a centrifuge. *J. comp. physiol. Psychol.*, 1951, **44,** 525–534.

Clark, B., Graybiel, A., & MacCorquodale, K. The illusory perception of movement caused by angular acceleration and by centrifugal force during flight: II. Visually perceived motion and displacement of a fixed target during turns. *J. exp. Psychol.*, 1948, **38,** 298–309.

Graybiel, A. Oculogravic illusion. *Arch. Ophthal., Chicago*, 1952, **48,** 605–615.

Graybiel, A., & Brown, R. H. The delay in visual reorientation following exposure to a change in direction of resultant force on a human centrifuge. *J. gen. Psychol.*, 1951, **45,** 143–150.

Graybiel, A., Clark, B., & Zarriello, J. J. Observations on human subjects living in a "slow rotation room" for periods of two days. *Arch. Neurol.*, 1960, **3,** 55–73.

MacCorquodale, K. Effects of angular acceleration and centrifugal force on nonvisual space orientation during flight. *J. aviat. Med.*, 1948, **19,** 146–157.

Noble, C. E. The perception of the vertical: III. The visual vertical as a function of centrifugal and gravitational forces. *J. exp. Psychol.*, 1949, **39,** 839–850.

Passey, G. E., & Guedry, F. E., Jr. The perception of the vertical: II. Adaptation effects in four planes. *J. exp. Psychol.*, 1949, **39,** 700–707.

Pearson, R. G., & Hauty, G. T. Role of postural experiences in proprioceptive perception of verticality. *J. exp. Psychol.*, 1960, **59,** 425–428.

Witkin, H. A. An investigation of problems of space orientation as related to the selection and training of pilots. *Off. Naval Res. tech. Rep.*, 1944. (Project No. N7-ONR-323).

(Received December 17, 1960)

Journal of Experimental Psychology
1962, Vol. 63, No. 3, 302–306

CONTEXT FACTORS IN PAIRED-ASSOCIATE LEARNING AND RECALL

DONALD M. SUNDLAND AND DELOS D. WICKENS

Ohio State University

The experiment reported in this paper deals with the effect of contextual cues as the discriminability of the primary stimuli is varied. Weiss and Margolius (1954) write, "It has been assumed by Guthrie, Hull, and McGeoch that responses can become associated not only with primary stimuli but also with context stimuli" (p. 318). In their experiment it was found that learning required fewer trials if there were differently colored backgrounds for the paired associates rather than uniform gray backgrounds. They also found that the recall, for groups learning under the colored backgrounds condition, was better when the color alone was shown than when the words were shown on uniform gray backgrounds. The context cues were more useful than the primary stimuli. The present experiment was designed to further clarify the role of context stimuli.

It was hypothesized that where the primary stimuli were highly discriminable, the addition of discriminable contexts would not facilitate the learning task. A corollary to this was that when the primary stimuli were less discriminable the addition to discriminable contexts would facilitate the learning task. The latter is merely a restatement of one of the results of Weiss and Margolius. It was also hypothesized that where the primary stimuli were highly discriminable, the recall to the associated discriminable contexts will be less than when the primary stimuli are poorly discriminable.

METHOD

Experiment I

Subjects.—The 114 *Ss* were female students, aged 18 to 20 yr., from an elementary psychology course, who chose to participate in this experiment to meet a requirement of the course. They were tested in 10 groups ranging in number from 8 to 16 *Ss*.

Apparatus.—A Patterson memory drum was fitted under an opaque projector. The image of the drum was projected onto a screen in a semidark room. The *Ss* were provided with booklets in which to write their responses. It consisted of lines on which the response could be written. There was a separate page for each trial. The lists were printed on white window-shade material in black India ink. The colors were painted on using "Permanent Pigments" water colors with a liquid detergent in the water, since otherwise the paint did not spread on the window-shade material. One set of primary stimuli were six-letter words selected from the first 1,000 most frequent words as listed in Thorndike (1944). The other set of primary stimuli were two-syllable, hyphenated, nonsense syllables, each chosen from the Glaze (1928) list of 7% association value. The responses were three-letter words selected from the first 100 most frequent words as listed in Thorndike (1944).

An attempt was made to avoid logical connections between the stimulus and the response and between the colored context and the response. The nine paired associates and the three rest interval pairs (TURN-NOW, NOW-REST, START-NOW) were printed on a circular tape with five replications, the order of the paired associates being randomized. Table 1 presents the responses, the stimuli, and the colored contexts.

Procedure.—The *Ss* were told that they were to attempt to anticipate the appropriate response by writing it in their booklets and that a new page would be used for each trial. After Trials 5, 10, and 15 the stimuli only were shown. These are called the test trials. Following each test trial a 30-sec. rest interval was introduced, during this time *E* circulated

through the room and looked at the booklets to check whether *Ss* had understood the instructions. An *S* was discarded if he responded to the rest interval pairs, or if *S* had written down all of the responses from the second learning trial on since it is unlikely that *S* could have learned the list in one trial and this would indicate he had misunderstood the instructions. The stimulus was shown alone for 2 sec. and the S-R pair was then shown for 2 sec. Two groups were tested under each of five learning conditions: color only (C); words only (W); nonsense syllables only (NS); words on colored backgrounds (C+W); one nonsense syllables on colored backgrounds (C+NS). After the regular test following Trial 15, the C+W and C + NS groups had two additional test trials. One group under each of these two learning conditions was tested first with the C list and then with the W (NS) list. The other group was tested in the reverse order.[1]

TABLE 1

STIMULUS AND RESPONSE MATERIALS

Response	Stimulus Items		
	Words	Nonsense Syllables	Colors
ARM	SPIRIT	ZIH-MIF	GREEN
ACT	FLOWER	FUH-YIG	UMBER
COW	SUDDEN	ZUB-YEK	ROSE
BAG	MANNER	QAS-ZEG	GREEN
OAK	DINNER	DAQ-GUK	VIOLET
WAY	BRANCH	BOF-NUX	YELLOW
CUP	DEGREE	CEJ-KEB	RED
DOG	MINUTE	MIB-VUF	ORANGE
OIL	LENGTH	GEC-ZOS	BLUE

Experiment II

This experiment was undertaken to obtain additional data which might facilitate an understanding of the factors which were operating to produce the results obtained in the first experiment. It was conducted in precisely the same manner and with the same materials and type of *Ss* as in Exp. I, but included only two major groups of 28 *Ss* each, color plus words, and color plus nonsense materials. In the retention test half the *Ss* were tested to the color first and half to the symbol first. Following the last retention condition they were asked to describe the nature of the stimulus they had employed during the learning. Four possible choices were given to *Ss*: symbols only; colors only; symbols and colors together; symbols sometimes and colors sometimes. The *Ss* who chose any stimulus using colors were requested to write down the names of the colors they used.

[1] We are grateful to Gloria Barts for her help in preparing the experimental materials and in collecting the experimental data. The research was supported in part by NSF Grant G6203 to the junior author.

TABLE 2

MEAN CORRECT RESPONSES ON TEST TRIALS 5, 10, AND 15: EXP. I

Cond.	*N*	Test Trial					
		5		10		15	
		Mean	*SD*	Mean	*SD*	Mean	*SD*
C	30	3.50	2.23	5.21	2.21	6.87	1.86
W	18	4.67	3.06	7.05	1.81	7.72	1.72
NS	24	2.96	2.19	6.13	2.01	6.88	1.88
C + W	18	4.16	2.52	6.94	1.87	8.33	0.95
C + NS	24	2.83	2.19	5.88	2.01	6.96	1.88

RESULTS

Experiment I

Tables 2 and 3 summarize the results of this experiment. The index of learning used was the number of correct responses on the test trials. The degree of recall was indicated by the number recalled correctly divided by the number of correct responses on the last regular anticipation trial, this ratio being expressed as a percentage.

TABLE 3

PERCENTAGE RECALL TO ELEMENT: EXP. I

Recall Order	*N*	Symbol		Color	
		Mean	*SD*	Mean	*SD*
C+W→W→C	12	100	11.6	6.7	10.8
C+W→C→W	12	95.5	8.2	14.9	13.8
C+NS→NS→C	12	77.8	31.6	37.8	34.5
C+NS→C→NS	12	67.8	19.5	56.7	31.9

TABLE 4

ANALYSES OF VARIANCE OF RECALL SCORES ON TEST TRIALS 5, 10, AND 15: EXP. I

Source	*df*	Trial 5 *F*	Trial 10 *F*	Trial 15 *F*
Between *Ss*	3			
Context (C)	1	.24	.16	.57
Stimuli (S)	1	6.79*	4.72*	2.62
C × S	1	.10	.00	4.90*
Error	80			
Total	83			

* $P < .05$.

That *Ss* were able to learn the list under Cond. C proves that the colored backgrounds were discriminable and sufficient for the learning task. Ten of 30 *Ss* in Cond. C had all correct on the last test trial.

Since the groups varied in number, *Ss* were discarded randomly until the cells of a Context × Stimuli design contained proportionate numbers. Table 4 shows the results of separate analyses of the three learning test trials. At no trial did the addition of colored contexts effect the learning performance. For Trials 5 and 10 Groups W and C + W attained higher scores than Groups NS and C + NS. At Trial 15 there was a significant interaction between the context and stimulus variables. However, this was in the unpredicted direction, Group C + W behaving as if the colored contexts added to their performance.

TABLE 5

PERCENTAGE RECALL TO AN ELEMENT: EXP. II

Recall Order	*N*	Symbol		Color	
		Mean	*SD*	Mean	*SD*
C+W→W→C	14	101.4	12.3	10.3	16.0
C+W→C→W	14	81.6	35.7	8.7	15.7
C+NS→NS→C	14	73.5	33.3	39.3	34.9
C+NS→C→NS	14	82.8	37.4	36.5	35.9

The second hypothesis implied that the use of the context cues alone would be more disturbing to the word group than to the nonsense group. The data was analyzed by means of a 2 × 2 × 2 analysis of variance. The columns and rows of the design corresponded, respectively, to the learning stimulus conditions (C + W vs. C + NS) and the test conditions (context alone vs. stimulus alone). The third dimension of the design represented the order of testing. Of the three main effects, only the test conditions (color or symbol) affected the degree of retention, *Ss* showing more retention to the symbol ($F = 115.01$, $df = 1/88$, $P < .01$). Of the four interactions, only S × T was significant ($F = 34.25$, $df = 1/88$, $P < .01$). As hypothesized, Group C + NS was able to give more correct responses when presented with the colored contexts alone than Group C + W. The recall percentage for the word groups to the context cue alone was only 10.8% as opposed to 47.8% for the nonsense groups ($P < .001$). To the symbol alone, however, the word groups excelled

TABLE 6

PERCENTAGE RECALL TO AN ELEMENT ACCORDING TO THE STIMULUS USED: EXP. II

Route	*N*	Symbol		Color	
		Mean	*SD*	Mean	*SD*
C + W	28				
W only	26	90.6	29.7	6.8	12.6
Sometimes	2	100	.00	44.5	11.5
C + NS	28				
NS only	10	100.4	12.1	1.7	5.1
NS and C both	6	89.5	16.3	54.3	12.9
Sometimes	8	77.3	27.3	49.9	11.1
C only	4	5.5	5.5	79.0	18.5

the nonsense groups, with 97.8% and 72.8%, respectively ($P = .001$).

Experiment II

The results for the recall conditions are presented in Table 5. An analyses of variance gave essentially the same results as were obtained in Exp. I in that: (*a*) *Ss* showed a higher degree of retention to the symbols than to the colors ($F = 111.86$, $df = 1/104$, $P < .01$); (*b*) Group C + NS showed a higher degree of retention to colors than did C + W; (*c*) Group C + W showed a higher degree of retention to the symbols than did C + NS (for Stimulus × Test Element interaction, $F = 13.18$, $df = 1/104$, $P < .01$).

To probe more deeply into the factors which resulted in these retention differences, the groups were divided on the basis of *Ss*' stated stimulus preferences during learning. The data are shown in Table 6. It would seem that for Group C + W the vast majority of the *Ss* reacted to the symbols. However, in Group C + NS the situation was quite different and the majority of the *Ss* either supplemented the symbols with the colors or they used the colors alone.

For the 18 *Ss* of Group C + NS who used color in some way, a rank order correlation was run between the number of colors named and the percentage recall to color. The magnitude of the rho was .84 ($P < .001$).

Discussion

The results seem to indicate that the effect of context cues is a relatively complex one. There is no evidence in the data that the context cue facilitates learning, regardless as to whether the symbolic cue is a meaningful word or a nonsense syllable. This is true despite the fact that the context cue alone was as adequate a stimulus for learning as the nonsense syllables—that is, nonsense syllable alone, color alone, and nonsense syllable plus color are essentially the same. The results for words with and without color follow a similar pattern. When, however, the symbolic cue is removed for those who have learned with both present, the word group does extremely poorly; the nonsense group does not deteriorate to so marked an extent. Removing the context cue, however, has little, if any, effect upon the word group although it seems to depress the nonsense group slightly.

The results for the word group do not seem to be difficult to understand. The context cue seems to be essentially irrelevant to their behavior; it does not facilitate their learning, and it is essentially useless to them in the absence of the word. Apparently the highly discriminable verbal symbol dominates the stimulus situation during learning and suppresses the potential effectiveness of the context cue. It is as if these *Ss* failed to respond to the context cue as a potentially helpful stimulus during the learning. The verbal reports obtained in Exp. II do indeed indicate that the vast majority of these *Ss* reported using only the verbal symbols as stimuli.

When the symbolic stimulus is a nonsense syllable, however, the context cue acquires sufficient habit strength so that it alone can serve as a fairly effective stimulus in the retention-transfer situation. It might be expected that the presence of this additional and potentially effective dimension of differences among the stimuli would have facilitated performance during the learning of Group NS + C, but there is no evidence that such is the case. The verbal reports suggest an explanation of this result. Only 6 *Ss* stated that they used both dimensions of the stimulus, the others reporting that they used one or the other alone, or sometimes one and sometimes the other. In short, most of them behave as if only a single dimension were available at each S-R pair. The learning data offer some slight but nonsignificant evidence that the readiness to use both

or either dimension may have had a facilitating effect, for when the color or nonsense syllable alone *Ss* were combined and contrasted with the both or sometimes *Ss*, the mean for the dual *Ss* on Trial 15 was 8.07 as compared with 7.57 for the single *Ss*. Such a result, if real, could arise from differences in ability levels of *Ss* who are alert enough to attempt to make use of both dimensions as contrasted with those who chose only one, as well as to a difference in the efficiency of the two methods. A different type of experimental design from that employed in the present experiment will be required to answer this question as well as to discover the conditions which are more favorable to the utilization of both dimensions of the stimulus.

Weiss and Margolius (1954) had found that context cues alone were more effective response guides than were the symbolic cues when they were slightly modified. In the present experiment the context cues proved to be almost completely ineffective as sole cues in the situation where they had been combined with words, but fairly effective when employed with nonsense syllables. In both their experiment and in the present one the context cues alone could be used as stimuli for the learning.

The present experiment does not in any essential way contradict the Weiss and Margolius (1954) study, but rather can be combined with it to extend our knowledge concerning the operation of context cues. Briefly, the experiments suggest that effectiveness of context cues will be a function of an interaction between the discriminability of the major cues among themselves and the discriminability of the context cues among themselves. This experiment has only shown that context cues will have greater or lesser cue value as a function of the nature of the symbolic stimulus, but it is plausible to assume that the same relationship would hold if context discriminability were varied and symbolic cues held constant.

Summary

An experiment was conducted investigating the learning of paired associates, with and without differently colored backgrounds, and with "high frequency" words and low association value nonsense syllables as stimuli. The various lists were learned by 114 *Ss* in groups, using a memory drum attached to an opaque projector.

It was found that the addition of differently colored backgrounds did not increase the ease of learning either where the stimuli were high frequency words or where they were low association value nonsense syllables. The associations between the colored backgrounds and the responses were better reported when the nonsense syllables were the original primary stimuli than when the words were. The experiment suggests that effectiveness of a context cue will be a function of an interaction between the discriminability among the individual primary cues and the discriminability among the context cues.

A subsequent experiment was conducted in which *Ss* were asked to describe the stimuli they used in learning the list. Their verbal reports suggested that the results could be attributed, at least in part, to a tendency to select only one of the two available cues.

References

Glaze, J. A. The association value of nonsense syllables. *J. genet. Psychol.*, 1928, **35**, 255–269.

Thorndike, E. L. *The teachers workbook of 30,000 words.* New York: Teachers College, Columbia University, 1944.

Weiss, W., & Margolius, G. The effect on context stimuli on learning and retention. *J. exp. Psychol.*, 1954, **48**, 318–322.

(Received December 23, 1960)

Journal of Experimental Psychology
1962, Vol. 63, No. 3, 307–313

TESTS OF THREE HYPOTHESES ABOUT THE EFFECTIVE STIMULUS IN SERIAL LEARNING[1]

ROBERT K. YOUNG

University of Texas

The items of a serial task are presented in a constant order; because of this the effective stimulus for a specific response may possibly be one of several alternatives: It may be: (*a*) the item immediately preceding the response to be learned; (*b*) a combination of several items preceding the response to be learned; (*c*) the position which the item holds in the list; (*d*) some combination of these; or (*e*) something else entirely.

The present paper reports a series of three experiments designed to help identify the effective stimulus in serial verbal learning. The first alternative (*a*), called the specificity hypothesis (Young, 1961), is tested in Exp. I by measuring transfer from a serial list to a subsequent paired-associate (PA) list composed of pairs taken from adjacent items in the serial list. If associations are formed between adjacent items in serial learning, positive transfer should obtain from the serial to the PA list and the PA transfer pairs should be learned faster than control pairs which have not been previously learned in serial order.

The second alternative (*b*), called the compound-stimulus hypothesis, is tested in Exp. II by measuring transfer from a serial list to a PA list composed of items taken from the serial list. Positive transfer from the serial to the PA task would be taken as evidence favorable to a compound-stimulus interpretation of the effective stimulus in serial learning.

The third alternative (*c*), called the serial-position hypothesis, is tested in Exp. III by observing transfer from one serial list to a second composed of the same items as the first. The serial position hypothesis assumes that items holding the same serial position in both lists (S items) would show positive transfer from the first to the second list while items holding different positions in both lists (D items) would show negative transfer from the first to the second list. The reason for this is that S items constitute an A-B, A-B paradigm of positive transfer in which the stimulus (A) would be the constant serial position, and the response (B) would be the same item from list to list. The D items would constitute an A-B, A-Br paradigm of negative transfer (Porter & Duncan, 1953) in which the response to be learned to a constant position stimulus would be rearranged from the first list to the second.

General Method

In the three experiments all *Ss* were selected from introductory psychology classes at the University of Texas. All lists were constructed of dissimilar adjectives taken from Melton (Hilgard, 1951), and all were learned by the anticipation method. The lists were displayed by a filmstrip projector. The PA lists were learned at a standard 2:2-sec. rate with a 4-sec. intertrial interval while the serial lists were learned at a standard 2-sec. rate with a 6-sec. intertrial interval. When serial and PA lists were learned in an experiment, a 5-min. interval was employed

[1] The research reported in the present paper was supported in part by the National Science Foundation (G-4850) and in part by a grant from the Graduate School of the University of Texas.

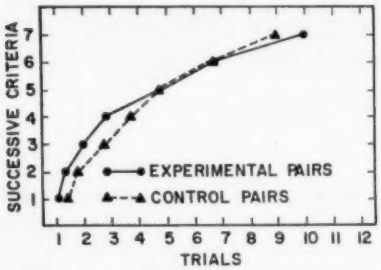

FIG. 1. Trials to successive criteria curves for the experimental and control sublists during PA learning in Exp. I.

between lists to allow time to change film strips and to read PA instructions; and when two serial lists were learned, a 1-min. interval between lists was needed to change film strips. All lists (with the exception of Exp. Ia) were learned to one errorless recitation and all PA lists had four orders of presentation to reduce the possibility of serial learning.

EXPERIMENT I

Method.—Each *S* learned a serial list and then a PA list. One-half the items in the PA list were experimental pairs obtained from adjacent items in the previously learned serial list, and the other half of the PA items were control pairs. The ordering of the experimental pairs in the PA list was exactly the same as in the serial list. That is, for these pairs the same stimulus elicited the same response in both serial and PA learning.

To reduce the possibility of differences due to intrinsic differences in difficulty in the pairs themselves, the items which served as experimental pairs for half the *S*s served as control pairs for the other half. Subgroups of 18 *S*s learned one of two serial lists and then the common PA list containing items from both serial lists (Total $N = 36$). Two 14-item serial lists (plus a cue symbol) were constructed and one 14-pair PA list was derived from the two lists.

Results.—The PA list may be thought of as being composed of two seven-pair sublists. The mean trials to learn the experimental and control sublists to one errorless recitation were 10.06 and 9.06, respectively ($F = .97$, $df = 1/34$). The mean correct responses per *S* during learning for the experimental and control sublists were 62.58 and 63.22, respectively ($F = .17$, $df = 1/34$).

As in a previous experiment measuring transfer from a serial to a PA list (Young, 1959), positive transfer occurred early in learning and then disappeared later in learning. The mean number of trials to attain successive criteria are plotted in Fig. 1. The *F*s obtained between the experimental and control sublists for criteria of one, two, three, and four correct responses ($df = 1/34$ in each case) were 4.86 ($P = .05$), 5.25 ($P = .05$), 10.50 ($P = .01$), and 7.61 ($P = .01$), respectively. Beyond a criterion of four correct responses, the two sublists do not differ with respect to mean trials to successive criteria.

The associations of the experimental pairs were differentially reinforced during serial learning. For example, Item 2 in the serial list was correct .84 times per trial; in contrast, Item 10 was given .39 times

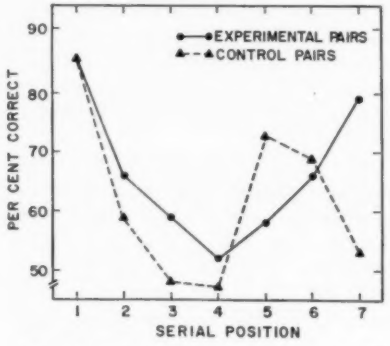

FIG. 2. Mean percentage correct in PA learning for the experimental and control sublists in Exp. I with the responses being ordered along the baseline as they were in the serial lists.

per trial. This difference reflects, of course, the bowed serial-position curve commonly found in serial learning. Since adjacent pairs were taken from the serial list and used in the PA learning, serial position effects should transfer to the PA learning if associations are formed between adjacent items. Where the pairs in the two PA sublists were arranged according to position in their respective serial lists, the mean percentage frequencies with which these experimental and control pairs were correct during PA learning are shown in Fig. 2. Of the seven comparisons made, one significant difference was obtained. As can be seen in Fig. 2, the last pair in the experimental sublist was given correctly more often than its control ($F = 28.73$, $df = 1/34$, $P < .001$).

The PA error data were analyzed in two ways. The stimuli in the PA list can be classified either as experimental or as control, and the data were analyzed according to this classification. The mean numbers of errors per trial elicited by experimental and control stimuli were .52 and .39, respectively. This difference is significant ($F = 5.87$, $df = 1/34$, $P < .05$). Secondly, the mean numbers of incorrect experimental items per trial elicited by an experimental stimulus and the mean numbers of control items per trial elicited by a control stimulus were .44 and .32, respectively ($F = 7.05$, $df = 1/34$, $P < .05$). It is noteworthy that the remainder of the errors per trial, .08 and .07, is small and constant across conditions.

Experiment Ia

The associations formed during serial learning may not have been of sufficient strength to transfer to the subsequent PA task. Experiment Ia was conducted to control for this possibility. The design of this experiment is exactly the same as Exp. I with one exception: the serial learning criterion was increased to one errorless recitation plus an additional 10 trials. A total of 22 *Ss* were employed with 11 per subgroup.

The mean numbers of correct responses per *S* given to the even items during serial learning, arranged in serial position, may be summarized as follows: 36.36, 27.09, 25.36, 19.59, 22.55, 24.68, and 31.77. These even items were the response members of the seven experimental pairs in the subsequent PA learning. If it can be assumed that number of correct responses is a measure of associative strength, then the present experiment permits the formation of associations between sucessive items in the serial list.

The mean trials to learn the experimental and control PA sublists to one errorless recitation were 8.86 and 8.14, respectively ($F = .85$, $df = 1/20$). The mean correct responses per *S* during learning for the experimental and control sublists were 49.59 and 51.91, respectively ($F = .84$, $df = 1/20$). As in Exp. I, transfer of serial position effects, transfer during early learning, and overt errors were analyzed. In none of these analyses was a significant difference obtained.

Experiment II

Method.—The only difference between Exp. II and Exp. I was in the construction of the PA list. Two 15-item serial lists (plus a cue symbol) were constructed and a 10-item PA list was constructed from the two serial lists. Each stimulus in the PA list was composed of two items. These items were arranged one over the other and were in the same sequence (from top to bottom) as they were in the serial list. Forty-six *Ss* served in the experiment with 23 learning each serial list before learning the PA list.

Results.—As in Exp. I, the PA list may be thought of as being composed of two five-pair sublists. The mean numbers of trials required to reach one errorless recitation of the experimental and control sublists were 12.04 and 9.63, respectively ($F = 8.35$, $df = 1/44$, $P < .01$). As can be seen from the data on trials to successive criteria presented in Fig. 3, negative transfer obtained from the serial to the PA list.

The mean numbers of correct re-

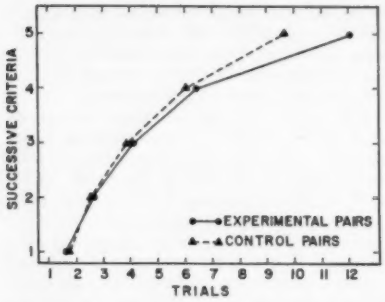

FIG. 3. Trials to successive criteria curves for the experimental and control sublists during PA learning in Exp. II.

sponses per *S* given during learning of the experimental and control sublists were 41.13 and 46.91, respectively ($F=7.27$, $df=1/44$, $P<.01$). Both response measures are in agreement that prior serial learning retards subsequent PA learning.

As in Exp. I, the pairs in the two PA sublists were arranged according to position in their respective serial lists and the mean percentage frequencies correct for the experimental and control pairs are presented in Fig. 4. Only one of the five comparisons was significant. The second experimental pair was more difficult than was the comparable control pair ($F = 20.49$, $df = 1/44$, $P < .001$).

The results of the error analysis in Exp. II are comparable to the error data obtained in Exp. I. Experimental stimuli elicit more overt errors than do control stimuli ($F = 19.70$, $df = 1/44$, $P < .001$) with mean errors per trial being .61 and .34 for the experimental and control sublists, respectively. In addition, more experimental items are elicited incorrectly by experimental stimuli (.56 per trial) than is the case for control stimuli eliciting control items (.29 per trial) ($F = 23.91$, $P < .001$). Again it should be noted that the remainder of the errors per trial, .05 and .05, is small and constant across conditions.

EXPERIMENT III

Method.—Experiment III differs from Exp. I and II in that transfer from one serial list to a second was studied and, in addition, a control group was employed. Four 13-item serial lists were constructed (with the first item serving as a cue symbol) for use in the experiment. Lists A and B were experimental lists and List C was a control list. Each *S* learned one of these lists and then all *S*s learned the test list. Comparing List A with the test list, the even items in List A held the same serial position as they did in the test list while the odd items were randomly rearranged from List A to the test list. In List B the odd items held the same serial position as they did in the test list while the even items were randomly rearranged. In this manner test-list differences between those items with the same serial positions (S items) and those items with different serial positions (D items) are attributable to transfer effects rather than to differential difficulty due either to specific items or to serial position. The Control List C had no items in common with the test list, and test-list learning in the two experimental conditions may be compared to that in the control condition to evaluate transfer effects. Sixty-

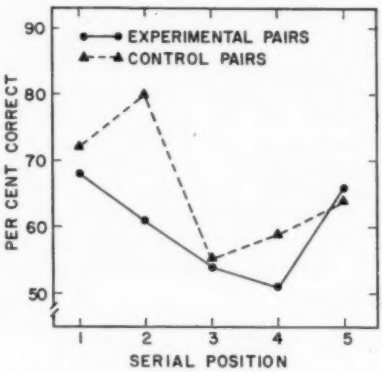

FIG. 4. Mean percentage correct in PA learning for the experimental and control sublists in Exp. II with the responsess being ordered along the baseline as they were in the serial lists.

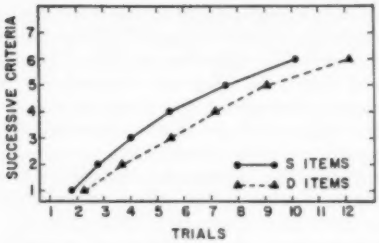

FIG. 5. Trials to successive criteria curves for the S- and D-item sublists during test-list learning in Exp. III.

three *Ss* served in the experiment, 21 *Ss* in each condition.

Results.—The mean numbers of trials to learn the test list in Conditions A, B, and C were 12.43, 13.86, and 9.71, respectively ($F = 6.50$, $df = 2/60$, $P < .01$). Orthogonal comparisons (Senders, 1958) between the three conditions indicate that the experimental conditions did not differ ($F = 1.50$) while the test list in the control condition was learned faster than in the experimental conditions ($F = 11.51$, $df = 1/60$, $P < .01$). Negative transfer obtained from the experimental to the test lists and these data are comparable to transfer data obtained by Irion (1946).

The test list may be thought of as being composed of two six-item sublists—the S items and the D items. Trials to reach a criterion of one errorless recitation for each of these sublists in the experimental conditions were analyzed. The mean trials to learn the S-item sublist was 10.17, while the mean for the D-item sublist was 12.14 ($F = 14.97$, $df = 1/40$, $P < .001$). The mean numbers of trials to attain successive criteria are plotted in Fig. 5.

The mean numbers of correct responses per *S* during test-list learning for the S- and D-item sublists were 46.86 and 39.90, respectively ($F = 35.24$, $df = 1/40$, $P < .001$). Both response measures are in agreement that the S-item sublist was easier to learn than the D-item sublist.

The S and D items were arranged according to their respective serial positions in the test list and, as can be seen from Fig. 6, differences obtained as a function of serial position. Only the first and last items did not differ between the S-item and D-item sublists. The *F*s for Serial Positions 2, 3, 4, and 5 ($df = 1/40$ in each case) were 6.82 ($P = .05$), 6.54 ($P = .05$), 10.59 ($P = .01$), and 6.79 ($P = .05$), respectively. Thus the items in the middle of the list were learned faster when the serial position was retained than when it was changed while such differences were not observed for the items at the ends of the list.

DISCUSSION

Experiment I tested the hypothesis that associations are formed between successive items within a serial list. This hypothesis, called the specificity hypothesis, was tested by constructing pairs taken from successive items in a serial list and measuring transfer from the serial list to a PA list composed of

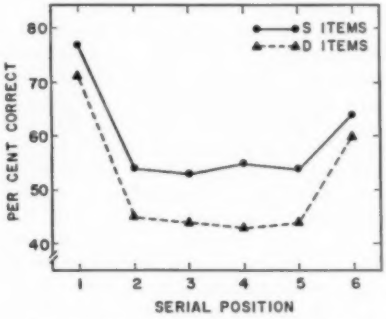

FIG. 6. Mean percentage correct in test-list learning for the S- and D-item sublists with the items being ordered along the baseline as they were in the test list.

the serial pairs. The analyses of the data indicated that, aside from positive transfer early in learning, the experimental pairs were learned no faster than the control pairs.

It may be argued that the absence of an overall difference between the two sublists was due to interference resulting from the change of tasks. However, when the arrangement of the items between serial and PA task is such as to expect negative transfer, zero transfer also occurs (Young, 1961).

On the other hand, the serial-position hypothesis assumes that the position the item holds in the list is the effective stimulus rather than the item immediately preceding the response to be learned. If such were the case, no transfer would be expected from the serial to the PA list because the stimulus in the PA learning would be different from the effective stimulus in serial learning. Thus, according to the serial-position hypothesis, the transfer paradigm employed in Exp. I would be an A-B, C-B paradigm of zero transfer. In this case the A stimulus would correspond to the serial position of the item in the serial list, the C stimulus would correspond to the nominal stimulus in the PA list, and the B response would be the response to be learned in both lists. Thus, while the specificity hypothesis predicts positive transfer from the serial to the PA list, the serial-position hypothesis predicts zero transfer.

Positive transfer occurred early in learning in Exp. I. This could be taken as evidence favoring a specificity notion, or it could be accounted for in terms of gross differentiation. Such an interpretation would state that following serial learning *S* was able to differentiate between those items which were learned previously and those which were not. When presented with a prefamiliarized item, *S* would restrict his responses to some other prefamiliarized item. Some of these guesses would be correct (significantly more than the control early in learning) and some would be incorrect (significantly more than the control errors).

The hypothesis that the two items preceding the response to be learned is the effective stimulus in serial learning was tested in Exp. II. Contrary to expectations from the compound-stimulus hypothesis, negative transfer occurred. While these data could be explained in terms of interference between the stimulus items (specificity interpretation) or in terms of interference arising from the similarity of the PA stimulus items to the initial serial positions (serial-position interpretation), the results of Exp. II do not permit serious consideration of the compound-stimulus hypothesis as a useful explanation of the stimulus in serial learning at this time.

The hypothesis tested in Exp. III was that the position the item holds in the serial list is the effective stimulus in serial learning. This hypothesis was tested by observing transfer from one serial list to another. The same items were employed in both lists and were arranged so that S items retained the same serial position in both lists while D items were randomly rearranged from the first to the second list. If the serial position were the effective stimulus, positive transfer would be expected for the S items and negative transfer would be expected for the D items. The results of the experiment show that the S items were easier to learn than the D items and that the test list in the experimental conditions was learned slower than in the control. The specificity hypothesis would make the prediction that no differences would obtain between the S and D items in the test list. According to this position, each item was associated to a response in the first list; but in no case were the same two items learned successively in both lists. Thus, the transfer paradigm for all items becomes an A-B, A-Br paradigm of negative transfer.

The addition of a mediation construct would permit interpretation of the results of Exp. III in terms of the specificity hypothesis. The only requirement would be the assumption that during test-list learning an S item acts as a stimulus and elicits the next (missing)

item, which in turn correctly elicits the next (S) item. Such a mediation assumption would enable the specificity hypothesis to predict that the S items would be learned faster than the D items.

The data of the present experiments provide no final answer to the question of the effective stimulus in serial learning. They are, for example, difficult to integrate into the extensive literature on remote associations (e.g., McGeoch, 1936) which may best be interpreted in terms of the specificity hypothesis. However, the serial position curves of Exp. I and III (Fig. 2 and 6) may be used to formulate a tentative hypothesis which employs both the specificity and the serial-position hypotheses. Evidence from Exp. I and elsewhere (Young, 1961) indicates that learning in accordance with the specificity hypothesis occurs at the extremes of the serial list but not in the middle. In addition, data from Exp. III indicate that position learning has an important cue function in the middle of the list but not at the extremes. It then may be that specificity learning occurs at the extremes and that position learning occurs in the middle of the list. Some difficulties are immediately apparent with this formulation: the middle of the list is the point of greatest "position confusion" (Schulz, 1955), i.e., position learning should be less likely to occur; and when degree of serial learning is increased as in Exp. Ia, evidence for specificity learning is not apparent.

Summary

Three hypotheses regarding the effective stimulus in serial learning were tested. If associations are formed between items of a serial list, then these associations should transfer to a subsequent PA list composed of the serial items with the S-R arrangement of the PA list being consistent with the ordering of the items in the serial learning. Positive transfer was observed only early in learning. If the effective stimulus in serial learning is the two preceding items in the list, then when sets of two items are used as stimuli, positive transfer should be observed in a subsequent PA task. In this case, negative transfer occurred. Finally, if the position of the item in a serial list is the effective stimulus, positive transfer should obtain from one serial list to another when the item retains the same serial position; and negative transfer should obtain when the serial position is changed. It was found that items retaining the same serial position were learned faster than items which had their serial positions changed from one list to another.

References

Hilgard, E. R. Methods and procedures in the study of learning. In S. S. Stevens (Ed.), *Handbook of experimental psychology.* New York: Wiley, 1951. Pp. 517–567.

Irion, A. L. Retroactive inhibition as a function of the relative serial position of the original and interpolated items. *J. exp. Psychol.*, 1946, **36**, 262–270.

McGeoch, J. A. The direction and extent of intra-serial associations at recall. *Amer. J. Psychol.*, 1936, **48**, 221–245.

Porter, L. W., & Duncan, C. P. Negative transfer in verbal learning. *J. exp. Psychol.*, 1953, **46**, 61–64.

Schulz, R. W. Generalization of serial position in rote serial learning. *J. exp. Psychol.*, 1955, **49**, 267–272.

Senders, V. L. *Measurement and statistics.* New York: Oxford Univer. Press, 1958.

Young, R. K. A comparison of two methods of learning serial associations. *Amer. J. Psychol.*, 1959, **72**, 554–559.

Young, R. K. The stimulus in serial verbal learning. *Amer. J. Psychol.*, 1961, **74**, 517–528.

(Received December 24, 1960)

Journal of Experimental Psychology
1962, Vol. 63, No. 3, 314–319

VERBAL CONCEPT ATTAINMENT: A FUNCTION OF THE NUMBER OF POSITIVE AND NEGATIVE INSTANCES PRESENTED

M. S. MAYZNER

University of Southern California

Kendler and his associates have suggested in a number of studies (Kendler & D'Amato, 1955; Kendler & Mayzner, 1956; Kendler & Vineberg, 1954) that the problem of concept attainment may be analyzed in terms of an implicit mediational mechanism. Most recently Kendler and Karasik (1958) have applied this approach to verbal concept attainment, and in the present study, we shall be concerned with developing this approach further.

In verbal concept attainment *S* typically is confronted with an array of instances or words. Usually, some subset of the total array is made up of those instances that *E* has designated as positive concept instances (i.e., words that belong to the concept), while the remaining subset is made up of those instances that *E* has designated as negative concept instances (i.e., words that do not belong to the concept). The *S*'s task is to identify or separate the positive from the negative instances, and to respond to the positive instances with the correct concept name.

According to a mediational hypothesis, *S* learns or attains the correct concept name in the following way. Each instance or word, whether positive or negative, evokes an implicit response. When more than one instance evokes the same implicit response, the probability of *S* grouping these different instances together increases, resulting in *S* explicitly responding to these different instances or words with the same concept name. If the instances selected and the associated concept corresponds to that subset of the total array that *E* designated as positive concept instances, then *S* is said to have attained the concept.

The employment of such a mediational model suggests a number of hypotheses concerning verbal concept behavior. First, it suggests that the ease or speed with which a given concept is attained is a function of the response strength of particular instances to evoke particular implicit responses. This hypothesis has been tested by Underwood and Richardson (1956b), who showed that the higher the dominance level (i.e., response strength), the more rapidly the concept was attained. Secondly, since implicit responses are evoked by all the instances, whether positive or negative, the type of implicit responses evoked by the negative instances (i.e., whether they are all different, all the same, have high or low response strength, etc.) may be as critically important as the implicit responses evoked by the positive instances. In fact, Kendler and Karasik (1958) have shown that different sets of negative instances, varying in the type of implicit responses they evoke, significantly affect the attainment of concepts based on a given set of positive instances. Thirdly, it might be expected that the actual number of positive and negative instances occurring in the total array should have a significant effect on concept attainment; that is, as the

TABLE 1

CONCEPTS, ASSOCIATED POSITIVE INSTANCES, AND ALL NEGATIVE INSTANCES USED IN THE STUDY

Test Problems

Concepts and Positive Instances								Negative Instances			
Round	%	Small	%	White	%	Hard	%				
BRACELET	25	POLLEN	30	HOSPITAL	32	ARMOR	24	ALUMINUM	CINNAMON	GINGER	OCEAN
MOON	30	THIMBLE	37	FROST	34	ASPHALT	29	APE	CLOVE	GYM	PINE
DIME	30	CABIN	39	BREAD	35	HELMET	31	BUTTER	COFFEE	HOG	SEAWEED
DERBY	33	SNAIL	42	PEARL	37	BRICK	35	CAMEL	CUSTARD	HONEY	SILK
PLATTER	38	WAIST	43	COLLAR	44	SKULL	36	CHEESE	FLANNEL	MOCCASIN	TAR
								CHOCOLATE	FOREST	MOSS	TOBACCO

Practice Problems

Concepts and Associated Positive and Negative Instances						
Red	%	Negative Instances	Sharp	%	Negative Instances
CHERRY	77	AMMONIA	FISHHOOK	70	CITY
TOMATO	83	BED	DAGGER	75	CORN
BEET	87	DANDELION	HATCHET	77	GARLIC
BLOOD	91	ELEPHANT	SCISSORS	78	GRASS
BLUSH	96	IVY	KNIFE	84	NIGHT
		VINEGAR			VELVET

number of positive instances, at a given level of response strength, is increased, there should occur a corresponding increase in the probability of more than one of them evoking the same implicit response, which should produce more rapid concept attainment. Conversely, as the number of negative instances is increased, when each negative instance is designed to evoke a different implicit response, there is a corresponding increase in the total number of different implicit responses evoked, which might be expected to interfere somewhat with the evocation of a single implicit response to the positive instances, thus producing less rapid concept attainment. It is to this issue, that is the effect of the number of positive and negative instances occurring, that the present study is directed.

METHOD

Subjects.—The *S*s were 144 students from introductory psychology classes at New York University, 12 in each of 12 different experimental conditions.

Stimulus material.—The stimulus material used in the study was selected from a word-list prepared by Underwood and Richardson (1956a) which gives the response strength (i.e., percentage frequency) of each word to evoke a sense-impression response. Since each *S* was required to attain four concepts, namely, Round (R), Small (S), White (W), and Hard (H), four lists of five positive instances were selected from Underwood's list, and are presented in Table 1, along with the percentage frequency value with which each instance evokes its associated concept.

As can be seen in Table 1, the range of response strengths associated with each concept was kept as low and as nearly equal, as was possible, to provide a moderately difficult and approximately equivalent set of problems, which might be expected to most clearly reveal the effects of the variables under investigation. Also included in Table 1 are the 24 negative instances employed in the study. These negative instances were selected from Underwood's list on the basis of two criteria: (*a*) none possess any response strength to evoke any of the four problem concepts, at least as indicated by Underwood's norms; (*b*) the response strength to evoke other concepts was kept as low as possible, in order to

examine at a base level the effect of increased numbers of negative instances.

Procedure.—The present study is concerned with examining the effects of the number of positive and negative instances occurring, on the speed of concept attainment. To examine these effects, a 4 × 3 factorial analysis of variance design was selected. Thus, either two, three, four, or five positive instances were combined with either two, four, or six negative instances to yield 12 different experimental conditions, with 12 *S*s per condition. In order to minimize the possibility that the results might be a function of a specific concept problem, each *S* received four test problems involving the Concepts R, S, W, and H. Four different orders of presentation were employed, namely, RWSH, SRHW, WHRS, and HSWR, i.e., within each condition 3 *S*s received the order RWSH, another 3 *S*s received the order SRHW, etc.

The form in which each problem was presented was as follows. All instances, both positive and negative, were typed in lowercase letters on white 3 × 5 in. cards. Each problem array was placed on a table before *S*. At the start of each problem, *S* was told to look at the cards and *E* started a stopwatch. The problem array may best be understood from the instructions given to *S*, as follows:

> When I tell you to start the problem, you will notice before you two rows of cards with words typed on them. The top row, you will notice has four cards or words, while the bottom row has seven cards or words. There is a single word in the bottom row, which may be grouped with the four words in the top row, on the basis of a common sensory quality all five words share together. Your task is to point out which word in the bottom row may be grouped with the four words in the top row, and to indicate the basis of the grouping, i.e., tell me the common sensory quality on which you have based your selection.

These instructions apply to that experimental condition involving five positive instances and six negative instances, with analogous instructions tailored to fit the other 11 conditions, with respect to the number of positive and negative instances occurring. Thus, one positive instance always appeared in the bottom row of words, with either two, four, or six negative instances, and either one, two, three, or four positive instances always appeared in the top row of words.

The *S*s were assigned at random to the 12 experimental conditions. Each *S* received two practice problems and four test problems and the time to solution was taken as the dependent variable for study. In the two practice problems, the number of positive and negative instances was the same as the number of positive and negative instances presented in the four test problems. The practice problems, shown in Table 1, were made easier than the test problems to insure that *S* fully understood the nature of the problems and what was required of him to achieve a correct solution. The *S* was instructed that a correct solution involved both (*a*) choosing the correct instance in the bottom row of words, and (*b*) applying the correct concept name. Either one alone was not considered a correct solution and *S* was instructed to continue working on the problem when this occurred. If no correct solution was obtained within 2 min., *S* was given the correct solution, and the next problem was presented.

Although care was taken to insure that all positive instances belonging to a given concept were essentially equivalent (i.e., had approximately equal percentage frequency values) additional design precautions were taken in the construction of each of the four problems presented to *S*, as follows. In all conditions involving the use of five positive instances (four in the top row and one in the bottom row) each of the five positive instances occurred at least twice and two of the five occurred three times, in the bottom row, to insure that no one particular positive instance appeared relatively more or less frequently than any other positive instance, among the negative instances. Similar precautions were taken with the conditions involving four, three, and two positive instances and, in addition, the selection of the remaining positive instances for the top row (either three, two, or one) was designed to insure relatively equal frequency of use. The combinations or arrays of positive instances that were so constructed, were employed without change in the conditions involving two, four, or six negative instances. Further, the placement of the positive instance among the negative instances was designed to insure that the positive instance occurred with approximately equal frequency in each of the available positions in the row.

Precautions also were taken in the selection of the negative instances. In the conditions involving 6 negative instances, since each *S* received 4 problems, the entire pool of 24 negative instances was utilized, as follows. The 24 negative instances were

TABLE 2

MEDIAN SOLUTION TIMES (SEC.) AND NUMBER OF SOLUTIONS FOR EACH PROBLEM-POSITION AND EACH CONCEPT SUMMED ACROSS THE 12 EXPERIMENTAL CONDITIONS

Concepts	Measure	Problem-Position				
		1	2	3	4	Totals
Round	Median	35.0	25.0	18.5	20.5	22.0
	No. solving	24	24	27	27	102
Small	Median	120.0[a]	120.0	120.0	120.0	120.0
	No. solving	7	7	13	11	38
White	Median	77.5	57.0	25.5	56.5	56.0
	No. solving	21	21	26	22	90
Hard	Median	57.0	19.5	70.0	22.5	32.0
	No. solving	22	30	23	28	103
Totals	Median	97.0	62.0	41.5	41.5	
	No. solving	74	82	89	88	

Note.—N = 36 each cell.
[a] Time-limit = 120 sec.

typed on cards and shuffled, 6 cards were then drawn at random, then from the remaining pack of 18 cards, 6 more were drawn at random, and the remaining pack of 12 cards was then randomly separated into 2 packs of 6 cards each. The 4 separate packs were then randomly assigned to serve as the 6 negative instances, for each of the 4 problems presented to a given *S*. This procedure was repeated 12 times, to provide a different random order of negative instances for each of the 12 *S*s in a given experimental condition. The same 12 random orders were employed, however, in the conditions involving 2, 3, 4, and 5 positive instances. Analogous randomization procedures were applied to the conditions with 2 and 4 negative instances, but with only 2 or 4 cards being drawn at each drawing, in order to construct the separate problems. When the separate drawings, of either 2 or 4 cards, exhausted the total pool of 24 cards, the 24 cards were again shuffled and the drawing procedure was repeated.

RESULTS

The dependent variable selected for study was "time to solution" and since each *S* received four problems, it was intended to take the median of these four time scores for each *S*, as the measure for analysis of the main effects in the design. However, the problems were not of equal difficulty as shown by the results presented in Table 2, which gives the median solution times and the number of solutions attained within the 120-sec. time limit, for each problem-position and each concept summed across the 12 experimental conditions. These results indicate that the concept Small is a far more difficult problem than the other three concepts, which are approximately of equal difficulty. This conclusion is supported by a chi square analysis ($\chi^2 = 34.52$; $df = 3$; $P < .001$) contrasting the obtained number of solutions for all four concepts (R-102, S-38, W-90, and H-103) with the expected equal number. Further, there is no significant dif-

ference between Concepts R, W, and H ($\chi^2 = 1.07; df = 3; P > .70$). The effect of problem-position is clearest in the Totals column and the correlated t for the difference between Problem-Positions 1 and 4 was 2.15 ($df = 143$; $P < .05$), suggesting a significant practice effect for this type of problem.

Since the concept Small produced such a deviant distribution of time scores, the median of the three time scores on Concepts R, W, and H was selected as the criterion measure for use in the analysis of variance. Thus, there was made available for each *S* a single median time score, and means, *SD*s, and ranges of these median scores, for the 12 experimental conditions of the study, are presented in Table 3. The results in Table 3 indicate that, in general, concept attainment solution times increase with increased numbers of negative instances, but decrease with increased numbers of positive instances, as most clearly seen in the row and column means. An analysis of variance yielded *F* values of 6.33, 1.40, and .39, for the main effects associated with the number of positive instances, the number of negative instances, and the interaction term, respectively, with only the *F* associated with the number of positive instances reaching statistical significance ($df = 3/132$; $P < .001$). Bartlett's test yielded a chi square of 8.60 ($df = 11; P > .50$), indicating the analysis did not violate the assumption of homogeneity of variance.

Discussion

It was predicted, on the basis of an S-R mediational model, that increasing the number of positive instances should produce more rapid concept attainment, since each additional positive instance presented to *S* should increase the probability of more than one such positive instance evoking the same implicit response, and matching what *E* has designated as the correct concept. It also was predicted, that increasing the number of negative instances, when each negative instance is designed to evoke a different implicit response, should produce less rapid concept attainment, since the presence of many different implicit re-

TABLE 3

Concept Attainment Solution Times (Sec.) for Different Numbers of Positive and Negative Instances

Number of Negative Instances		Number of Positive Instances				
		2	3	4	5	Row Mean
2	Mean	68.92	57.00	38.25	30.25	48.61
	SD	51.59	39.71	39.58	31.83	
	Range	10–120	9–120	6–120	9–120	
4	Mean	77.08	42.08	57.17	36.75	53.27
	SD	42.30	37.27	46.56	24.92	
	Range	5–120	5–115	6–120	9–99	
6	Mean	90.00	57.25	62.42	41.25	62.73
	SD	32.08	46.30	43.54	39.53	
	Range	30–120	8–120	7–120	5–120	
Column Mean		78.66	52.11	52.61	36.08	

Note.—Based on medians for individual *S*s.

sponses might be expected to interfere with the evocation of a single implicit response to the positive instances. The results indicate that while the effects of both positive and negative instances are in the predicted direction, only the effect associated with the positive instances reaches statistical significance.

The failure of the negative instances to produce a significant effect might indicate that the negative instances do, in fact, have no effect on concept attainment. However, since three of the four columns in Table 3 (i.e., two, four, and five positive instances) show systematic increases in solution times as the number of negative instances are increased from two to six, it seems more reasonable to conclude that the negative instances produce the effect predicted, but that the effect is a weak one and the present design was not sufficiently sensitive to measure it.

One might ask under what conditions would an increase in negative instances be expected to produce a strong effect. In a mediational model, three principal factors determine the speed with which a concept is attained. They are: (*a*) the number of positive instances presented and the response strength associated with each positive instance to evoke the appropriate implicit response; (*b*) the number of negative instances presented and the response strength associated with each negative instance to evoke inappropriate or irrelevant implicit responses; and (*c*) any interaction which occurs between the implicit responses evoked by all the instances, both positive and negative (i.e., Kendler and Karasik's (1958) "response differentiation"). By systematically varying then the parameters of this model, one could determine the particular values or conditions necessary for the negative instances to produce a strong effect.

Summary

The present study examined the effects of the number of positive and negative instances on verbal concept attainment. Four values of positive instances (two, three, four, and five) were combined with three values of negative instances (two, four, and six) in a 4×3 factorial design with 12 *Ss* per condition. The results, predicted on the basis of an S-R mediational model, showed that increased numbers of positive instances produced significantly more rapid concept attainment, while increased numbers of negative instances produced, though not significantly so, less rapid concept attainment.

References

KENDLER, H. H., & D'AMATO, M. F. A comparison of reversal shifts and nonreversal shifts in human concept formation behavior. *J. exp. Psychol.*, 1955, **49**, 165–174.

KENDLER, H. H., & KARASIK, A. D. Concept formation as a function of competition between response produced cues. *J. exp. Psychol.*, 1958, **55**, 278–283.

KENDLER, H. H., & MAYZNER, M. S. Reversal and nonreversal shifts in card-sorting tests with two or four sorting categories. *J. exp. Psychol.*, 1956, **51**, 244–248.

KENDLER, H. H., & VINEBERG, R. The acquisition of compound concepts as a function of previous training. *J. exp. Psychol.*, 1954, **48**, 252–258.

UNDERWOOD, B. J., & RICHARDSON, J. Some verbal materials for the study of concept formation. *Psychol. Bull.*, 1956, **53**, 84–95. (a)

UNDERWOOD, B. J., & RICHARDSON, J. Verbal concept learning as a function of instructions and dominance level. *J. exp. Psychol.*, 1956, **51**, 229–238. (b)

(Received December 30, 1960)

Journal of Experimental Psychology
1962, Vol. 63, No. 3, 320

SUPPLEMENTARY REPORT: MONETARY INCENTIVE AND MOTIVATION IN DISCRIMINATION LEARNING—SEX DIFFERENCES

BETSY WORTH ESTES, LOUISE BRIGHTWELL MILLER,[1] AND MARY ELLEN CURTIN

University of Kentucky

Miller and Estes (1961) found that amount of monetary reward, even in a 50 to 1 ratio, failed to improve error scores in the acquisition of a visual discrimination by 9-yr-old boys. A detrimental reward effect was suggested, in that rewarded groups made significantly more errors than a knowledge-only (KOR) group. n Achievement (n Ach) was shown to be a significant variable, high scorers making fewer errors. This study provides similar comparisons for adult *Ss*.

Method.—The *Ss* were 48 female and 72 male students in introductory psychology classes. A 3 × 2 × 2 factorial design was used: (*a*) three incentive levels (knowledge of results [KOR], 1¢, and 25¢); (*b*) two n Ach levels (high and low); and (*c*) two sex levels. Design, apparatus, and procedure were the same as in the previous experiment, except that the high reward group received 25¢ per correct choice instead of 50¢, and the stimulus presentation interval was shortened to 1 sec. in order to increase the difficulty of the task for adults. Groups were matched on the basis of n Ach scores, father's occupation, and amount of financial support from parents.

Results.—The significant variables were sex (females made fewer errors) and the sex-incentive interaction, Table 1. Sex differences and interaction are shown in Fig. 1. The detrimental effect of the 1¢ reward for females is confirmed by *t* tests; there was no significant incentive effect among males.

Female superiority might reasonably be attributed to cultural differences between sexes in past experience and training since the task involved attention to minute facial detail. Although a detrimental effect with low reward was also found in the previous study, the parallel between the 9-yr-old boys and adult females ends here. With 50¢ reward, performance of the young boys was similar to that of the 1¢ group and hence was poorer than with KOR; performance of the adult females with higher reward (25¢) was *better* than that of the 1¢ group and similar to KOR.

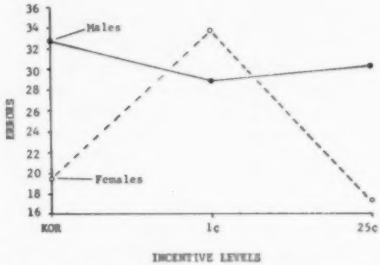

FIG. 1. Male and female performance according to incentive levels.

Age differences are shown not only by lack of incentive effect for adult males, but also by the n Ach factor, significant for young boys but not for male or female adults. One tenable explanation is that n Ach in adults was not measured by the method used. Another explanation involves the problem of reproduction of instructions when repeating with adults studies done with children. Atkinson (1954, p. 75) reports that high n Ach *Ss* perform better "only when the instructions defined performance as instrumental to producing feelings of accomplishment." Although instructions were experimentally achievement-oriented, it seems unlikely that the task of learning to tell drawings of twins apart could offer as much satisfaction in accomplishment for college students as for elementary school children, regardless of instructions.

[1] Now at the University of Louisville.

TABLE 1

ANALYSIS OF VARIANCE OF ERRORS

Source	*df*	*MS*	*F*
Sex (S)	1	1,433.69	5.30*
n Ach (N)	1	444.68	1.64
Incentive (I)	2	334.30	1.24
S × N	1	55.20	.20
S × I	2	1,025.17	3.79**
N × I	2	211.60	.78
S × I × N	2	444.32	1.64
Within cells	108	270.69	
Total	119		

* $P < .025$.
** $P < .05$.

REFERENCES

ATKINSON, J. W. Exploration using imaginative thought to assess the strength of human motives. *Nebraska symposium on Motivation: 1954.* Lincoln: Univer. Nebraska Press, 1954.

MILLER, L. B., & ESTES, B. W. Monetary reward and motivation in discrimination learning. *J. exp. Psychol.*, 1961, **61**, 501–504.

(Received March 23, 1961)

JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR

A New Monograph Series

Beginning with Volume 5 (1962), the *Journal of the Experimental Analysis of Behavior* will publish occasional monograph supplements. The first in the series is a research report by Jack D. Findley (University of Maryland), **"An Experimental Outline for Building and Exploring Multi-Operant Behavior Repertoires."** This appears as Part 2 of the January issue. The second monograph, **"A Review of Positive Conditioned Reinforcement,"** by Roger T. Kelleher and Lewis R. Gollub (Harvard Medical School and University of Maryland) is scheduled to appear late in 1962. These supplements will be sent to JEAB subscribers at no extra cost.

The regular subscription rate for libraries is $12.00 per annual volume. Both current and back volumes are still available at this rate, making it possible for librarians to acquire a complete file. For personal subscriptions, special rates are available for the current volume: $9.50 for members of recognized professional societies and $6.50 for students. All subscriptions begin with the January issue.

Please address all orders and business inquiries to:

KAY DINSMOOR
JEAB, Department of Psychology
Lindley Hall, Indiana University
Bloomington, Indiana

Checks should be made payable to the Journal